

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: October 12, 2004, 15:06:21 ; Search time 30 Seconds
(without alignments)
7158.527 Million cell updates/sec

Title: US-10-046-935-2234
Perfect score: 1962
Sequence: 1 atggacgtcgccggtgccc.....aatttgaatgaagcataa 1116

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_n2p_model -DEV=xlp
-O=/cgn2_1/USPTO_spool_p/US10046935/runat_12102004_155226_15255/app query.fasta_1.1287
-DB=PIR_79 -QFMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10046935 @CGN_1_1_46 @runat_12102004_155226_15255 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR_79.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					Description
Result No.	Score	Query Match	Length	ID	
1	602.5	30.7	202	T50635	hypothetical prote
2	289.5	14.8	462	D85438	hypothetical prote
3	250.5	12.8	515	F09700	protein F12A21.9 [
4	170	8.7	555	T01142	hypothetical prote
5	146	7.4	851	D86254	hypothetical prote
6	142	7.2	950	F86222	hypothetical prote
7	132.5	6.8	570	T49181	cyclophilin-like p
8	131.5	6.7	906	T01440	hypothetical prote
9	129	6.6	754	JCS314	CDG28/cdc2-like ki
10	117.5	6.0	595	T51747	RNA helicase RH26
11	115.5	5.9	1701	T43213	ENBP1 protein - ba
12	115	5.9	1051	G59436	KIAA1304 protein [
13	114	5.8	691	T48506	hypothetical prote
14	113	5.8	1641	T10955	early nodulin bind

15	111.5	5.7	1711	1	A47392	chromodomain-helic
16	107.5	5.5	395	2	T19144	hypothetical prote
17	107	5.5	561	2	T14792	hypothetical prote
18	106	5.4	447	2	T20552	hypothetical prote
19	105.5	5.4	857	2	T05352	hypothetical prote
20	105.5	5.4	1634	2	T26517	hypothetical prote
21	105	5.4	1200	2	T00042	BH-protocadherin p
22	105	5.4	1746	1	S19694	tenascin precursor
23	104.5	5.3	770	2	H84463	hypothetical prote
24	104.5	5.3	855	2	S56948	HAU5 protein - yea
25	104	5.3	519	2	S23796	modifier-3 protein
26	103.5	5.3	891	2	G84693	probable proline-r
27	103.5	5.3	1203	2	S26650	DNA-binding protei
28	103.5	5.3	1216	2	T09224	spindle assembly c
29	102.5	5.2	706	2	S43581	C28A5.6 protein (c
30	102.5	5.2	1061	2	G88428	protein C28A5.6 [i
31	102	5.2	560	2	T29586	hypothetical prote
32	102	5.2	925	2	T16235	hypothetical prote
33	101.5	5.2	451	2	T16418	hypothetical prote
34	101.5	5.2	932	2	T45894	hypothetical prote
35	101.5	5.2	1016	2	T25433	hypothetical prote
36	101.5	5.2	1269	2	T00443	hypothetical prote
37	101.5	5.2	1389	2	H84914	probable WD-40 rep
38	101	5.1	454	2	E84524	En/Spm-like transp
39	101	5.1	1252	2	T00263	hypothetical prote
40	101	5.1	1791	2	T02345	hypothetical prote
41	100.5	5.1	362	2	B64527	M-protein - Helico
42	100.5	5.1	553	2	S22954	finger protein zip
43	100.5	5.1	887	2	T39492	conserved hypothet
44	100.5	5.1	1289	2	T00387	KIAA0622 protein -
45	100.5	5.1	1307	2	G96711	unknown protein, 9

ALIGNMENTS

RESULT 1

T50635
hypothetical protein DKFZp762L0311.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 09-Jul-2004
C:Accession: T50635
R:Blum, H.; Bauersachs, S.; Newes, H.W.; Weil, B.; Wiemann, S.
submitted to the Protein Sequence Database, June 2000
A:Reference number: Z25145
A:Accession: T50635
A>Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-202 <AAA>
A:Cross-references: UNIPROT:Q9NPV2; EMBL:AL359619
A:Experimental source: adult melanoma (MeWo cell line); clone DKFZp762L0311
C:Genetics:
A>Note: DKFZp762L0311.1

Alignment Scores:
Pred. No.: 8.68e-44 Length: 202
Score: 602.50 Matches: 116
Percent Similarity: 69.23% Conservative: 28
Best Local Similarity: 55.77% Mismatches: 47
Query Match: 30.71% Indels: 17
DB: 2 Gaps: 4

US-10-046-935-2234 (1-1116) x T50635 (1-202)

QY	481	AGGCGTACATCCCG-----GGTTCCTCCAGGAGAAACCCCTGAACGGAGACTCGT	534
Db	5	ArgArgAlaPheSerGluGlyGlnIleThrArgArgMetAsnProThrArgSerAlaArg	24
QY	535	CCTCTTACCAAGTCAAGTCCCGGATCTCGGGTCCCTTGACGCTCTACCCATGGRGAGA	594
Db	25	ProProGluLysPhe-----AlaLeuGluAsnPhetheThrValSerAla	38
QY	595	GCCAGTTTKGAKMTMWTGTCATGTTGTGAGAAAGAGACCGTGGATGGCTACATG	654

Db 39 AlalysPheAlaGluGluPheTyrSerPheArgArgArgLysThrIleGlyGlyLysCys 58
QY 655 AATGAAGATGACCTGCCAGAGCGCTCCAGCATCATCGTGACCCCTCCGCATATA 714
Db 59 ArgGlu-----TyrArgArgArgHisArgIleSer-----Ser 69
QY 715 ATTGCCCGCAGTGAAGAAATTACAGAGAGGAGTGGAGAACGCTCTGCAGCAATTCTCGA 774
Db 70 PheArgProValGluAspIleThrGluGluAspLeuGluAsnValAlaIleThrAlaArg 89
QY 775 GAGAGATATATAACCGTCTCACTGGCTCTACTTGTCTCATCAATGCGCTCAGAGACTATT 834
Db 90 AspLysIleTyrAspLeuValLeuGlyAsnThrCysHisGlnCysArgGlnLysThrIle 109
QY 835 GATACCAAAACAACTGCAGAAACCCAGACTGCTGGCGGCTTCGAGGCCAGTCTGTGCGC 894
Db 110 AspThrLysThrValCysArgAsnGlnGlyCysCysGlyValArgGlyGlnPheCysGly 129
QY 895 CCCTGCTTCCGAAACCGTTATGTGTAAGAGGTGAGGATGCTCTGCTCGATCCGAACTGG 954
Db 130 ProCysLeuArgAsnArgTyrGlyGluAspValArgSerAlaLeuLeuAspProAspTrp 149
QY 955 CATTCGCCCGCTTGTCCAGGATCTGCAACTGCGAGTTTCTGCCGCGCAGAGATGCGCG 1014
Db 150 ValCysProProCysArgGlyIleCysAsnCysSerTyrCysArgLysArgAspGlyArg 169
QY 1015 TGTGCGACTGGGCTCTGTGTATTAGTAAATATCATGCTTTGGGAATGTGCATGCC 1074
Db 170 CysAlaThrGlyIleLeuIleHisLeuAlaLysPheTyrGlyTyrAspAsnValLysGlu 189
QY 1075 TACTTGAAGACCTGAAACAGGAA 1098
Db 190 TyrLeuGluSerLeuGlnLysGlu 197

RESULT 2
D85438
hypothetical protein AT4G37110 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: D85438
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: D85438
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-462 <STO>
A:Cross-references: UNIPROT:023175; GB:NC_001268; NID:g7270660; PIDN:CAB80377.1; GSPDB:G
C:Genetics:
A:Gene: AT4G37110
A:Map position: 4

Alignment Scores:	
Pred. No.:	9,95e-17
Score:	289.50
Percent Similarity:	38.32%
Best Local Similarity:	26.63%
Query Match:	14.76%
DB:	2
	Length: 462
	Matches: 98
	Conservative: 43
	Mismatches: 122
	Indels: 105
	Gaps: 16

US-10-046-935-2234 (1-1116) x D85438 (1-462)

QY 175 AGGACCCGCGCAGTGCAGGCGACTCTGACCTCTCAGGCTGGCGATGAAAGTTTCCAGCG 234
Db 5 ArgThrArgAlaLysCys-----SerIleProValThrAsnProAsn 18
QY 235 CGGAGTACAGGGGAGCAACCAACAAAGCAGAGTCCCGCCAGCCCTCAGAGAAATCT 294
Db 19 ProThrValGlyGly-----GlyAsnSerAsn 27
QY 295 GTGACTGATTCCAACTCCGATTCCAGAGATGAAGTGAATCAATTTTTCGAGAAAGG 354

Db 28 ValSerIleTyrGluLysCysArgGluAspArgIleLysGluAsnLeuGlnArgMetLys 47
QY 355 GCTTTAAATATAAGCAAAACAAAGCAATCTTGCAAAACCTCATGCTCTGAATAGAAAGC 414
Db 48 AsnLeuGlyIleMetAsp-----LeuSerLeuLysLeuLysSerGluIle----- 62
QY 415 TTCCCTGGCTGCTTCCCGTGGAGACATCCCTCCAGGCTCCGACTCAATCAAGAGGA 474
Db 63 -----Arg 63
QY 475 CCGGAAGGCGTACATCCCGGGTGTCTTCCAGGAGAAAACCTTGAAACGAGAGCTCGT 534
Db 64 ProAlaLysArgArgTyrGlyAsn-----SerAsnAlaAsnProGlyArgGluThrSer 81
QY 535 CCTCTT-----ACCAGGTCAAGGTCCCGATCCCTCGGCTCCCTT----- 573
Db 82 ProIleGlnLeuSerValSerSerArgArgSerSerArgValPheArgGluIleIleSer 101
QY 574 -----GACGCTCTACCCATGGRGAGAGCCAGTCTTGARKMTMWTCATGTTGGTG 624
Db 102 LeuPheTyrAsnGlnIleGluAlaGlyAlaIleSerPheSerGluIleLeuGluValIle 121
QY 625 AGA-----AAGAGAAGACCGTGGATGGCTACATGATGAAGATGACCTGCCAGAGAC 678
Db 122 SerSerLeuLysAsnAlaThrProValSerTyrAlaGluGluProGluLeuLysGly 141
QY 679 CGTGGCTCCAGATCATCCGTCACCTT-----CGCATATAATTCCGCCCTGGAAGAA 732
Db 142 LysValSerLysGluGluIleValLeuTrpValGlyGluGlyValArgPro-----Glu 159
QY 733 ATTACAGAGGAGGTGGAGAACGTC-----CGCATATAATTCCGCCCTGGAAGAA 759
Db 160 IleTyrThrGluGluHisGluLysLeuGlyAsnThrGluArgThrTrpGluLeuPhe 179
QY 760 -----TCGAGCAATCTCGAGAGAGATATATAACCGTTCACCTGGCTCTACTTGT 810
Db 180 ValAspGlyCysAspLysAsnGlyLysArgIleTyrAspProValArgGlyLysCysCys 199
QY 811 CATCAATGC-----CGTCAGAGAG 828
Db 200 HisGlnCysArgIleGluAlaAsnValPheAsnGlyTyrCysGlyPheTyrArgGlnLys 219
QY 829 ACTATTGATCAAAACAACTGCAGAAACCCAGACTGC-----TGGGGCGTTTCAGGCCAG 885
Db 220 ThrLeuGlyTyrHisThrGlnCys-----SerGlnCysAsnHisSerValArgGlyGln 237
QY 886 TTCTGTGGCCCTGCTTCGAAACCGTTATGGTGAGAGAGTCCAGGATGCTCTGCTGGAT 945
Db 238 PheCysGlyAspCysLeuTyrMetArgTyrGlyGluHisValLeuGluAlaLeuGluAsn 257
QY 946 CCGAACTGGCATTTGCCCGCTTGTCCAGGAAATCTGCAACTGTCAGTTTCTGCCGCGCAGGA 1005
Db 258 ProAspTrpIleCysProValCysArgAspIleCysAsnCysSerPheCysArgThrLys 277
QY 1006 GATGACCGGTGCGACTGGGGTCTT----- 1032
Db 278 LysGlyTyrLeuProThrGlyAlaAlaTyrArgLysIleGlnLysLeuMetAlaPhePro 297
QY 1033 -----GTGTATTAGCCAAATATCAT-----GGCTTTGGGAATGTGCATGCCCTACTTG 1080
Db 298 LeuAsnPhePheCysSerGlnIleHisLysLeuGlyTyrLysSerValAlaHisTyrLeu 317
QY 1081 AAAAGCCTGAAACAGGAAATTTGAA 1104
Db 318 IleGlnThrAsnGlnSerGlu 325

RESULT 3

F96700
protein F12A21.9 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: F96700
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Jensen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, C.A.; Salzborg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: F96700
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-515 <STO>

A;Cross-references: UNIPROT:Q9FXE1; GB:AE005173; NID:g11072030; PIDN:RAG28909.1; GSPDB:G

C;Genetics:
A;Gene: F12A21.9
A;Map position: 1

Alignment Scores:
Pred. No.: 2,34e-13 Length: 515
Score: 250.50 Matches: 48
Percent Similarity: 58.49% Conservative: 14
Best Local Similarity: 45.28% Mismatches: 37
Query Match: 12.77% Indels: 7
DB: 2 Gaps: 2

US-10-046-935-2234 (1-1116) x F96700 (1-515)

QY 778 AGATATATAACCGTTCACTGGGCTCTACTGTGATCAATCGCGTCAGAGACTATTGAT 837

Db 25 ArgTyrAspSerSerAsnGlyLysThrCysHisGlnCysArgGlnLysThrMetAsp 44

QY 838 ACCAAACAACATGC-----AGAAACCCAGACTGCTGGGCGGTTCCAGGCCAG 885

Db 45 PheValAlaSerCysLysAlaMetLysLysAspLysGlnCys-----ThrIleAsn 61

QY 886 TTCTGTGGCCCTCGCTTCGAAACCGTTATGTTGTAAGAGGTCAGGGATGCTCTGCTGGAT 945

Db 62 PheCysHisLysCysLeuIleAsnArgTyrGlyGluAsnAlaGluValAlaLysLeu 81

QY 946 CGAACTGGCATTCGCCGCTTCGAGGAATCTCAACTGCAATGCTGCGCGGACGCA 1005

Db 82 AspAspTyrLysCysProGlnCysArgGlyIleCysAsnCysSerPheCysArgLysLys 101

QY 1006 GATGGACGGTGCAGCTGGGCTCTGTGATTAGCCAAATATCATGCTTTGGGAAT 1065

Db 102 ArgGlyLeuAsnProThrGlyIleLeuAlaHisLysAlaLysAlaSerGlyLeuAlaSer 121

QY 1066 GTGCATGCTCTACTTGAAA 1083

Db 122 ValSerMetLeuLeuGlu 127

RESULT 4

T01142

hypothetical protein At2g23530 [imported] - Arabidopsis thaliana

N;Alternate names: hypothetical protein F26B6.18

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004

C;Accession: T01142; H84625

R;Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,

submitted to the EMBL Data Library, June 1998

A;Description: Arabidopsis thaliana chromosome II BAC F26B6 genomic sequence.

A;Reference number: Z14198

A;Accession: T01142

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-555 <ROU>

A;Cross-references: UNIPROT:O80469; EMBL:AC003040; NID:g3242700; PID:g3242716

A;Experimental source: cultivar Columbia

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: H84625

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-555 <STO>

A;Cross-references: GB:AE002093; NID:g3242716; PIDN:RAC23768.1; GSPDB:GN00139

C;Genetics:

A;Gene: F26B6.18; At2g23530

A;Map position: 2

A;Introns: 87/2; 166/2; 188/1; 231/2

C;Superfamily: Arabidopsis thaliana hypothetical protein F26B6.18

Alignment Scores:

Pred. No.: 2,1e-06 Length: 555
Score: 170.00 Matches: 78
Percent Similarity: 40.81% Conservative: 33
Best Local Similarity: 28.68% Mismatches: 92
Query Match: 8.66% Indels: 69
DB: 2 Gaps: 12

US-10-046-935-2234 (1-1116) x T01142 (1-555)

QY 274 CGCCAGCCCTCAGAGAAATCTGTGACTGATTCCAACTCCGATTTCAGAA-----GAT 324

Db 2 ArgThrGluAlaGlnAspSerValProLysSerAsnProAsnProGluLeuIleLysGlu 21

QY 325 GAAAGTGAATGAATTTTGGAGAAA---AGGGCTTTAAATAATAAGCAAAACAAAGCA 381

Db 22 ThrProLysValSerLeuTyrGluGlnCysArgGluGluArgLysGluAsnLeuGln 41

QY 382 ATGCTTGCAAACTCATCTCTGATTTAGAAAGCTTCCTGGCTCGTCCGTGGAAGACAT 441

Db 42 ArgMetAsnAsnLeu-----GlyLeuLeu 49

QY 442 CCCTCCCGAGGCTCCGACTCACAATCAAGAGACCCGAGAGGCTACATTCGCCGAGCCGT 501

Db 50 AsnLeuSerArgLysLeuLysProLysThrArgProValLysArgSerTyrGlyAsnArg 69

QY 502 GCTTCCAGAGAAACCTCGAAGAGAGCTGCTCTTTACAGGTCAAGGTCCCGGATC 561

Db 70 AsnSerValGlnAsnProThrProProLeuGlnProSerArgArgSer---SerArgLeu 88

QY 562 CTGGGTCTCTGACCTCTACCCATCGRGAGAGCCAGTTTGCARKMTWGTACATGTTG 621

Db 89 GluAsnThrThrProValIle----- 95

QY 622 GTGAGAAAGAGAGAGACCCGTGGATGGCTACATGAATGAAGATGACCTGCCAGAGCCGT 681

Db 96 -----TyrThrAspGly---IleAsnGluLysGlyLysLysAlaSerLys 109

QY 682 CGCTCCAGATCATCCGTGACCTTCGCGATATAATTCGCCAGTGGAGAAATTCAGAG 741

Db 110 -----ArgGluSerValValIleGlyGlyIleArg-----AlaGluIleTyrThr 125

QY 742 GAGGAGTTCGAGAACCTGCGACCAATTCGNA----- 774

Db 126 GluGluHisGluLysLeuLeuGlyAsnThrGluArgSerTyrThrCysPheValAspGly 145

QY 775 -----GAGAAGATATATAACCGTTCACTGGCTCTACTTGTCTATCATCATGC 819

Db 146 TyrAspLysAsnGlyLysArgIleTyrAspProPheAsnGlyLysThrCysHisGlnCys 165

QY 820 CGTCAGAACTATTGATACCAAAACAAACTGCAGAAACCCAGACTGTGGGCGGTCGA 879

Db 166 ArgGlnLysThrMetGlyHisArgThrGlnCys-----SerGluCysAsnLeuValGln 183

QY 880 GCCCAGTTCTGT-----GGCCCCCTCCCTCGAAACCGTTATGTCGAGAGGTCAGG 930

Db 184 GlyGlnPheCysAspMetValAsnMetCysLeuLys----- 195

QY 931 GATGCTCTCTGATCCGAACTGGCATTCGCCGCTTGTGCGAGGAATCTGCAACTGCACT 990

```

Db      196 ---LeuLeuArg-ileGlnThrGlyPheValProLeuValAlaGlySerAlaIleAlaVa 214
QY      991 TTCTCCGGCAGGAGATGGACGGTGTGGACTG 1024
Db      214 lCysAlaGlyIleIleLysAspGlyPheGlnLeu 225

RESULT 5
D86254
hypotheical protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: D86254
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Huizar, L.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: AB6141; MUID:21016719; PMID:11130712
A:Accession: D86254
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-851 <STO>
A:Cross-references: UNIPROT:065394; GB:AE005172; NID:g3157933; PIDN:AAC17616.1; GSPDB:GN
A:Genetics:
A:Map position: 1

Alignment Scores:
Pred. No.: 0.00261 Length: 851
Score: 146.00 Matches: 72
Percent Similarity: 40.40% Conservative: 50
Best Local Similarity: 23.84% Mismatches: 128
Query Match: 7.44% Indels: 52
DB: 2 Gaps: 17

US-10-046-935-2234 (1-1116) x D86254 (1-851)

QY      268 GAGTCCCGCAGCCCTCAGAG---AATCTGTGATGATTCGAACTCCGATTCGAGAAGAT 324
Db      28 GluLysLysLysProValGluAlaThrSerLeuSerSerGlySerSerAspIleGluGlu 47
QY      325 GAAAGTGGATGAATTTTGGAGAAAGCGCTTAAATATATAAGCAACAAACAAAGCAATG 384
Db      48 GluIleSerValGlu---CysProLysArgValAlaAsnGlnArgLysArgSerLys 66
QY      385 CTTGCAAACTCATGCTCAATTAGAAAGCTTCCCTGGCTCGTTCGTCGGAAGACATCCC 444
Db      67 AlaAspGluIleLysThrLys-----SerSerArgLysArgLysCys 80
QY      445 CTCACGAGCTCCGATCACAATCAAGAGACCGGAGCGGATACATCCCGGTTGTGCT 504
Db      81 AspAspGluAsnLysCysGluGluAsnGluLysLysGlnArgSer-----SerValLys 98
QY      505 TCCAGAGAAACCTCAACGGAGAGCTCGTCTCTTACCAGGTCAGGTCGCGATCCTC 564
Db      99 LysArgAlaThrThrTrpLysGluGluGluValValAspAspGluLysLysCysGlu 118
QY      565 GGGTCCCTTACGCTCTACCCATGGRGAGAGCGAGTTTKGARKMTWMTGATGTTGGTG 624
Db      119 GlnGlnLeuGlnLeuValProSerSerLysAlaThrSerArgSerArg-----Ser 135
QY      625 AGAAGAGAAACCGTGTGATGGTACATG-----AATGATGATGACCTGCCAGA---- 675
Db      136 LysLysSerValSerValAspThrTrpLeuValAsnAsnGlnIleAspValSerAlaLeu 155
QY      676 AGCGGTCCGATCATCCGTGACCCCTCCGCATATAATTCGCCCGCATGAAGAAATT 735
Db      156 SerSerArgSerGluSerGluLeuSer-----AspSerTyr 167

```

```

QY      736 ACAGAGGAGGAGTTGTCAGAACGTCCTGACAGCAATTTCTCGAGAGAGATATATAACCGTTCA 795
Db      168 LeuLysThrGluTyrPheAsnAspCys---ArgSerMetThrArgSerLeuLysAlaAsn 186
QY      796 CTGGGC-----TCTACTTGTTCATCAATGC-----CGTCAGAGACTATTGTGATACC 840
Db      187 LeuGlyGluLeuAlaIleCysHisGlnCysSerLysGlyGluArgArgTyrLeuPheIle 206
QY      841 AAACAAACTGCAGRAACCCAGACTGCTGGGGCGTTTCGAGGCCAGCTTCTGTGGCCCTCGC 900
Db      207 CysThrPheCysGlu-----ValArg-----LeuTyrCysPheProCys 219
QY      901 CTTCAAAACCGTATGTGTGAAGAGGTCAAGGATGCTCTGCTGCATCCGAACATCGCATTCG 960
Db      220 IleLysLysTyrTyrProHisLeuSerThrAspAspIleLeuGlu-----LysCys 236
QY      961 CGCCCTTGTGAGGAATTCGCACTGCACTTCTCCGGCAGCGAGATGAGCGTGTGGC 1020
Db      237 ProPheCysArgGlyThrCysAsnCysCysThrCysLeuHisSerSerGlyLeuIleGlu 256
QY      1021 ACTGGGGTCTTGTGTATTATTAGCCAAATATCATGCTTTGGGAATGTGCAT----- 1071
Db      257 ThrSerLysArg---LysLeuAspLysTyrGluArgPheTyrHisLeuArgPheLeuIle 275
QY      1072 -----GCCTACTTCAAAAGCCTG-----AAACAGGAATTTGAAATG 1107
Db      276 ValAlaMetLeuProPheLeuLysLysLeuCysLysAlaGlnAspGlnGluLeuThr 295
QY      1108 CAAGCA 1113
Db      296 GluAla 297

RESULT 6
F86222
Hypotheical protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: F86222
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Huizar, L.
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: AB6141; MUID:21016719; PMID:11130712
A:Accession: F86222
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-950 <STO>
A:Cross-references: UNIPROT:004024; GB:AE005172; NID:g2342679; PIDN:AAB70402.1; GSPDB:GN
A:Genetics:
A:Map position: 1

Alignment Scores:
Pred. No.: 0.000585 Length: 950
Score: 142.00 Matches: 78
Percent Similarity: 32.15% Conservative: 31
Best Local Similarity: 23.01% Mismatches: 128
Query Match: 7.24% Indels: 102
DB: 2 Gaps: 15

US-10-046-935-2234 (1-1116) x F86222 (1-950)

QY      103 GATGACAGTTGTGACAGCTTTCCTTCTGTGATAATTTTGCACACGAGGTCAGTCAGTT 162
Db      6 AspGluThrCysAspSerVal----- 12
QY      163 CGGAGAGGCTGTAGNCCCGCAGCCAGTCAGGCATCTTGGACCTCTCAGGTTGGCGATG 222

```



```
Db 13 -----:::|::|::|
QY 223 AGTTTCCAGCGGAGTACCAGGGAGAGCAACCAACAAAAGCAGAGTCCCGCCAGCCC 282
Db 18 AsnGluGlnThrArgSerAlaAsnGlyIleGlyAsnGlyAsnGlyGluSerIleProGly 37
QY 283 TCAGAGAATTCTGTGACTGATTCACAACTCCGATTTCAGAGAT----- 324
Db 38 IleProAspLeuArgCysLysArgSerAspGlyLysGlnTrpArgCysThrAlaMet 57
QY 325 -----GAAAGTGGGAATGAATTTTGGAGAAAAGGGCTTTA 360
Db 58 SerMetAlaAspLysThrValCysGluLysHisTyrIleGlnAlaLysLysArgAlaAla 77
QY 361 AAT-----ATAAGCAAAACAAAGCAATGCTTGCAAACTCATGCTGAATTA 408
Db 78 AsnSerAlaPheArgAlaAsnGlnLysLysAlaLysArgSerSerLeuGlyGluThr 97
QY 409 GAAAGCTTCCCTGGCTCGTTCCGTGGAAGACATCCCTCCCA----- 450
Db 98 AspThrTyrSerGluGlyLysMetAspPheGluLeuProValThrSerIleAspHis 117
QY 451 -----GCTCCGACTCACAATCAAGGAGACCGCGAAGCGGTACA 489
Db 118 TyrAsnAsnGlyLeuAlaSerAlaSerLysSerAsnGlyArgLeuGluLysArgHisAsn 137
QY 490 TTCCCGGGTGTCTTCCAGGAGAACCTCGACGAGAGCTCGTCTCTTACCAGTCA 549
Db 138 -----LysSerLeuMetArgTyrSerProGlu-----ThrProMetMetArgSer 152
QY 550 AGGTCC---CGGATCCTCCGGTCCCTT---GACGCTCTACCCATGGRGAGAGCCAGTTTK 603
Db 153 PheSerProArgValAlaValAspLeuAsnAspAspLeuGlyArgAspValValMetPhe 172
QY 604 GARKMTWGTACATGTGTGTGAGAAAGAGAGACCGTGGATGGCTACATGAATGAAGAT 663
Db 173 GluGluGlyTyrArgSerTyrArgThrProProSerVal---AlaValMetAsp----- 189
QY 664 GACCTGCCAGAGCGCTCGCTCCAGATCATCGTGACCTCCGATATATTCGCCCA 723
Db 190 -----ProThrArgAsnArgSerHisGlnSerThrSer-----Pro 201
QY 724 GTGGAAGAAATTACAGAGGAGGAGTTGGAGAACGCTCTGCAGCAATCTCGAGAGAAGATA 783
Db 202 MetGlu-----TyrSerAlaAlaSerThrAspVal 211
QY 784 TATAACCGTTCACTGGCTCTACTTGTCTATCATGTCGTCAGAAAGACTATTGATACCAA 843
Db 212 SerAlaGluSerLeuGlyGluIleCysHisGlnCysGlnArgLysAspArgGluArgIle 231
QY 844 ACAAACTGCAGAAACCCAGACTCTCTGGGGCGTTCGAGGCGAGTTCGTGGCCCTCCCTT 903
Db 232 IleSerCysLeuLysCysAsn-----GlnArgAlaPheCysHisAsnCysLeu 247
QY 904 CGAAACCGTTATGGT-----GAAGAGTTCAGGGAGTCTCTGTGGATCCGAAC 951
Db 248 SerAlaArgTyrSerGluIleSerLeuGluValGluLysVal----- 262
QY 952 TGGCATTCGCGCTTGTGAGGAATCTGCAACTGCACTGAGTTTCTGCGGAGCGAGAT 1008
Db 263 -----CysProAlaCysArgGlyLeuCysAspCysLysSerCysLeuArgSerAsp 279
```

RESULT 7

```
T49181
cyclophilin-like protein - Arabidopsis thaliana
N;Alternate names: protein MAA21.30
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C;Accession: T49181
R;Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Rudd, S.; Lemcke, K
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z25018
```

```
A;Accession: T49181
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-570 <XRI>
A;Cross-references: UNIPROT:O9LY75; EMBL:AL163818; GSPDB:GN00061; ATSP:MAA21.30
A;Experimental source: cultivar Columbia; BAC clone MAA21
C;Genetics:
A;Gene: ATSP:MAA21.30
A;Map position: 3
A;Introns: 26/3; 45/1; 74/3; 83/1; 96/1; 135/2; 189/1; 341/2; 455/2; 482/2
F;6-175/Domain: cyclophilin homology <CYP>
```

Alignment Scores:

Pred. No.:	0.00363	Length:	570
Score:	132.50	Matches:	79
Percent Similarity:	39.09%	Conservative:	41
Best Local Similarity:	25.73%	Mismatches:	108
Query Match:	6.75%	Indels:	80
DB:	2	Gaps:	15

US-10-046-935-2234 (1-1116) x T49181 (1-570)

```
QY 4 GAGCGTCCGCGCGTCCGAGAAAGATCTCAGAGTAAGAAAGAACTTAAAGAAATTCAGA 63
Db 207 GluAlaLysGluThrArgLysLysGluSerAsnGluLysArgLysArgLysArg 226
QY 64 TATGTCAAGTTGATTTCATGGAACCTCGTCATCTCTGATGCACAGTTGTGCACGC--- 120
Db 227 TyrSerSerSerAspSerTyrSerSerSerSerSerSerSerSerSerSerSerGlu 246
QY 121 ---TTGCTTTCGATAATTTTGCAAACACGAGCTGCAGTTCGCGAAGGCTGTAGG 177
Db 247 AlaTyrSerSerSerSerTyrGluSerSerSerSerSerSerSerSerSerSerSer 266
QY 178 ACCCGCAGCAGTCGAGGCACCTCTGACCTCTCAGGGTGGCGATGAAGTTTCAGCGCG 237
Db 267 ---LysSerThrThrArgHisLysGly-----ArgArgGlyGluArg 279
QY 238 AGTACCAGGGGACCAACCAACAAA---GCNAGCTCCGCGCAGCCCTCAGAGAT 291
Db 280 LysSerLysGlyArgSerGlyLysLysLysAlaArgProAspArgLysProSerThrAsn 299
QY 292 TCTGTG-----ACTGATTCCCAATCCGATTCAGAAAGATGAAAGTGAATGAATTTT 345
Db 300 SerSerSerAspThrGluSerSerSerSerSerSerSerSerSerSerSerSerSer 315
QY 346 GAGAAAAGGGCTTTAAATAAAGCAAAACAAAGCAATGCTTGCAAAACTCATGTCTGAA 405
Db 316 GlyHisLysAla-----IleLysSerValLysValAspAsnAlaAspGlnHisAlaAsn 333
QY 406 TTAGAAAGCTTCCCTGGCTCGTTCGTTGGAAGACATCCCTCCAGGCTCCGACTCACAA 465
Db 334 LeuAspAsp-----SerValLysSerArg 341
QY 466 TCAAGGAGACCCGAGGCGGTACATTCGCGGTGTGTCTTCAGGAGAAACCCCTGAACCG 525
Db 342 SerArgSerProIleArg-----ArgArgAsnGlnAsnSer 353
QY 526 AGAGCTGCTCTCTTACCAAGGTCAAGGTCCCGATCTCCGGTCCCTTGCAGCTCTACCC 585
Db 354 ArgSerLysSerProSerArgSerProValArgValLeuGlyAsnGlyAsnArgSerPro 373
QY 586 ATGGRGAGCCAGTTTKGARKMTWGTACATGTTGGTCAGAAAGAGNAGACCTGGAT 645
Db 374 -----SerArgSerProValArg 379
QY 646 GGCTACATGAATGA-----AGATGACTGCCAGAAAGCGTCCGCTCCAGTATCATCGTGAC 701
Db 380 AspLeuGlyAsnGlySerArgSerProArgGluLysProThrGluGluThrValGlyLys 399
QY 702 CTTCCGCATATAATTCGCCCTAGTGA-----AGAAATTAACAGAGGAGGTGGAGAA 755
Db 400 SerPheArgSerProSerProSerGlyValProLysArgIleArgLysGlyArgGly--- 418
```

QY 756 CGTCTGCAGCAATCTCTCAGAGAGAGATATATAACGGTTTCACTGGGCTCTACTTGTGTCATCA 815
Db 419 -----PheThr----- 420
QY 816 ATGCGTCAGAGACTATTGTATACCAAAACAACTGCAGAAACCCAGACTGCTGGGGCGT 875
Db 421 -----GluArgTyr-SerPheAlaArgLysTyrHisThrProSer---ProGluArg 436
QY 876 TCGAGGCCAGTCTCTGTGGCC 896
Db 437 SerProProArgHisTrpPro 443
RESULT 8
T01440
hypotheical protein F2401.2 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C/Accession: T01440
R/Shinn, P.; Buehler, E.; Dewar, K.; Feng, J.; Kim, C.; Li, Y.; Sun, H.; Conway, A.; Con-
eologis, A.; Ecker, J.R.
submitted to the EMBL Data Library, January 1998
A:Description: Genomic sequence for Arabidopsis thaliana BAC F2401.
A:Reference number: Z14211
A:Accession: T01440
A>Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-906 <SHI>
A:CROSS-references: UNIPROT:048794; EMBL:AC003113; MID:g2689438; PID:g2781345; GSPDB:GNO
A:Gene: ATSP:F2401.2
A:Map position: 1
A:Introns: 239/2; 272/3; 317/1; 338/2; 682/3; 832/1; 860/3
Alignment Scores:
Pred. No.: 0.00469 Length: 906
Score: 131.50 Matches: 70
Percent Similarity: 33.87% Conservative: 36
Best Local Similarity: 22.36% Mismatches: 90
Query Match: 6.70% Indels: 117
Gaps: 14
US-10-046-935-2234 (1-1116) x T01440 (1-906)
QY 253 ACCAACAAAAAGCAGAGTCCCGCAGCC-----TCAGAGAATTTCTGTGACTGATCC 306
Db 22 ThrThrArgLysLysArgThrLeuLysProPheMetSerLysGlySerSerProSer 41
QY 307 AACTCCGATTC-----GAAGATGAAGTGAATG 336
Db 42 SerSerAspSerArgLysArgLysLeuSerArgAlaGluAspSerAspSerAlaVal 61
QY 337 AATTTTTCGAAAAAGGCTTTAATATA----- 366
Db 62 LysArgAsnAlaLysArgArgLysLysCysLysValGluGluTyrTyrGluAspAsp 81
QY 367 -----AAGCAAAAC 375
Db 82 AspCysIleLeuSerAspTrpValGlnArgAsnThrAlaLysArgIleAspLysArgAsn 101
QY 376 AAGCAATGCTTGCAAACTCATGTCTGAATTAGAAGCTTCCCTGGCTCGTCCGTGA 435
Db 102 GluGluVal-----GluValMetValLysIleGluSer----- 112
QY 436 AGACATCCCTCCAGGCTCCGACTCACAATCAAGGAGACCGGAAGGGGTACATCCCG 495
Db 113 -----GlyAspAspCysThrIle-----GlyLysTrpPheSer 123
QY 496 GGTGTGCTTCAGGAGAAACCTCGAACGAGAGCTCGT----- 534
Db 124 AspValSerSerLysArgLysAspLysArgGlnValGluValAspGluAspGluTrp 143
QY 535 -----CCTCTTACCAGTCAAGGTCCTCCGATCCTC 564

Db 144 GluGluGluValThrLeuCysSerLysIleLysAlaThrSerSerArgSerArg---Thr 162
QY 555 GGGTCCCTTGCAGCTCTACCCATGGRGAGACCCAGCTTTTGARKMTWGTACATGTTGGTG 624
Db 163 HisSerLeuSerAla----- 167
QY 625 AGAAAGAGGAAGACCGTGGATGGCTTACATGAATGAAGATGACCTGCCAGAAAGCCGCTGC 684
Db 168 -----AsnSerProGluAsnValThrAspValIleSerProCysArg 181
QY 685 TCCAGATCATCCGTGACCCCTTCCGCATATAATTCGCCAGTGGAGAAATAATTCAGAGGAG 744
Db 182 SerArgSer-----ProAlaSerAsnValSerAspSer 192
QY 745 GAGTTGGAGAGCTCTGCAGCAATTTCTCAGAGAGAGATATATAACCGTTCTACTGGGCTCT 804
Db 193 IleGlnLysAsnAspCysThrSerArg-----LysGlnSerGlyPro 207
QY 805 ACTTGTCATCAATCCGTCAGAGACTATTGATACCAAAACAAACTGCAGAAACCCAGAC 864
Db 208 IleCysHisGlnCysLeuLysGlyGluArgIleThrLeuLeuLeuCys-----SerGlu 225
QY 865 TGCTGGGGCGTTCGAGGCCAGTTCTGTGGCCCTGCTCGAAACCCGTTATGTTGAAGAG 924
Db 226 Cys-----GluLysThrMetPheCysLeuGlnCysIleArgLysTrpTyrProAsnLeu 243
QY 925 GTCAGGATGCTCTGCTGATCCGAACTGCGCATTCGCCCTTGTTCGAGGAATCTGCAAC 984
Db 244 SerGluAspValValGlu-----LysCysProLeuCysArgGlnAsnCysAsn 260
QY 985 TGCAGTTCTTCGCGCAGCAGATGACGCTGTGCGACT 1023
Db 261 CysSerLysCysLeuHisLeuAsnGlyLeuIleGluThr 273
RESULT 9
JC5314
CDC28/cdc2-like kinase associating arginine-serine cyclophilin - human
N:Alternate names: CARS-Cyp
C:Species: Homo sapiens (man)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C/Accession: JC5314; G02262
R:Neatel, F.P.; Colwill, K.; Harper, S.; Pawson, T.; Anderson, S.K.
Gene 180, 151-155, 1996
A:Title: RS cyclophilins: Identification of an NK-TR1-related cyclophilin.
A:Reference number: JC5314; MUID:97128820; PMID:8973360
A:Accession: JC5314
A:Molecule type: mRNA
A:Residues: 1-754 <NES>
A:CROSS-references: UNIPROT:Q13427; EMBL:U40763; NID:g1117967; PIDN:AAB40347.1; PID:g111
A:Experimental source: thymus
A:Note: submitted to the EMBL Data Library, November 1995
C:Comment: This protein contains a cyclophilin-related domain, two 140K nuclear phosphop
splicing by binding to splicing factors containing serine-arginine repeats protein.
C:Genetics:
A:Gene: GDB:CYP; CARS-CYP
A:CROSS-references: GDB:9956062
C:Superfamily: CARS cyclophilin; cyclophilin homology
F,7-177/Domain: cyclophilin homology <CYP>
Alignment Scores:
Pred. No.: 0.00753 Length: 754
Score: 129.00 Matches: 76
Percent Similarity: 38.48% Conservative: 51
Best Local Similarity: 23.03% Mismatches: 145
Query Match: 6.57% Indels: 58
Gaps: 12
US-10-046-935-2234 (1-1116) x JC5314 (1-754)
QY 4 GACGCTCGCGGTCGCGCAGAAAGATCTCAGAGTAAGAGAACTTAAAGAAATTCAGA 63
Db 437 AspIleArgArgAsnSerGluLysAspLysTyrLysAsnLysValLysArgAla 456

[illegible]

RESULT 10

RESOL
T51747

RNA helicase RH26 [imported] - *Arabidopsis thaliana* (fragment)

C:Species: Arabidopsis thaliana (mouse-ear cress)

```
C:\projects\magazops\data (maggie car cross)
C:\Date: 18-Aug-2000 #sequence revision 18-Aug-2000 #text change 09-Jul-2004
```

C:Accession: T51747

R:Aubourg, S.: Kreis, M.: Lecharny, A.

Nucleic Acids Res. 27, 628-636. 1999

A;Title: The DEAD box RNA helicase family in Arabidopsis thaliana.

A:Reference number: Z22965; MUID:9862990; PMID:9862990

A:Accession: T51747

A; Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-595 <AUB>

A;Residues: 1-333 (A0B)
A:Cross-references: UNIPROT:O9ZRZ9; EMBL:AJ010474; PIDN:CAA09213.1

A:Experimental source: cultivar Columbia
A;CROSS-References: ONTFR01:Q9ZKZ3; EMBL

A; Experiment
C: Genetics:

A:Gene: RH26

A:Gene: RH26
A:Map position: 5

A;map position: 5
C:Superfamily: ATP-dependent RNA helicase DBP1

Alignment Scores:	
Pred. No.:	0.0719
Score:	117.50
Percent Similarity:	35.26%
Best Local Similarity:	23.40%
Query Match:	5.99%
DB:	2
Length:	595
Matches:	73
Conservative:	37
Mismatches:	106
Indels:	15
Gaps:	15

US-10-046-935-2234 (1-1116) x T51747 (1-595)

QY	6	CGTCGCGCGCTGCGCGCAGAAAGATCTCAGAGTAAGAAGAACATTAAAGAAATTCAGATA	65
DB	41	ArgGlyProArgArgGlyArgGlyGlnSerAspArg	53
QY	66	TGTGAAGTTGATTTCCATGGAAACCTCGTCATCTCTGATGACAGTTGTGACAGCTTTGC	125
DB	54	-----PheGlyGly-AlaLysArgGlyLysGluGlyGluMetAspArgPheGln	69
QY	126	TTCT-----GATAATTTTGGAAAAACACGAGGCT	152
DB	69	yserProAsnArgArgArgThrSerGlyGluProAlaAspSerPheGlyAsnLysArgLue	89
QY	153	GCAGTCAGTTCCGGAAGGCTGTAGACCCCGCAGCCAGTGCAGGCACCTCTGGACCTCTCAG	212
DB	89	uGlyAsp---ArgGluGlySerArgAsnGly-----Ar	99
QY	213	GGTGGCGATGAAGTTTCAGCGCGGAGTACCAGGGGAGCAACCAACAAAAAGCAGAGTC	272
DB	99	gValGlnGlyLysSerSerGluSerSerPheArgGlyArgSerAspArgAsnValAspSe	119
QY	273	C-----CGCCAGCCCTCAGAGAAATCTGTGACTGATTCCACTCC-----	312
DB	119	rGlySerSerPheArgGlyArgSerAspLysAsnValAspSerGlySerSerPheArgGln	139
QY	313	-----CATTCAGAAGATGAAGTGGAAATCAATTTTGGAGAAAGGGCTTTAAATAT	365
DB	139	yArgAsnAspArgAsnValAspSerGlySerSerPheArgGlyArgSerAspArgAsnVal	159
QY	366	AAAGCAAAACAAAGCAATGCTTTGCAGAACTCATGTCTGAAATAGAAAGCTTCCCTGGCTC	425
DB	159	lAspSerGlySerPheArgGlyArgSerAspArgAsnValAspSer--GlySerSe	178
QY	426	GTTCGGTGAAGACATCCCTC-----CCAGGCTCCGACTCAATCAAGGAGACC	476
DB	178	rPheArgGlyArgAsnAspArgAsnValAspSerGlySerSerPheArgGlyArgAsnAs	198
QY	477	CGCAAG-----CGTACATTTCCGGGTGTGCTTCCAGGAGAAACCCCTGA	521
DB	198	pArgAsnValGluSerGlyPheArgArgGluProGlySerGluAsnAsnArgGlyLeuGln	218
QY	522	ACGAGAGCTCGTCTCTTACC-----AGTCAAGGTCCCG	557
DB	218	ylYsglnThrArgGlyLeuSerLeuGluGluAspSerSerAspAspAspGluAsnAr	238
QY	558	GATC---CTCGGGTCCCTTGACGCTTACCCATGGGAGAGCCAGT-----TTKGA	605
DB	238	gValGlyLeuGlyAsnIleAspAspLeuProSerGluAspSerSerAspGluAspAspGln	258
QY	606	RKMTWGTACATGTTGGTGAAAGAGG-----	633

Db 258 uAsnAspGluProLeuLeuLeuLysLysAlaAlaSerAlaLysAlaValGlnThrAspLysPr 278
 QY 634 -----AAGACCGTGGATCGGTACATGAATGAA-----GATGACCT 668
 Db 278 oThrGlyGluHisValLysThrSerAspSerYrLeuSerLysThrArgPheAspGlnPh 298
 QY 669 GCCAGAACCGTGGCTCCAGATACCTCGT----- 699
 Db 298 eProLeuSerProLeuSerLeuLysAlaIleLysAspAlaGlyPheGluThrMetThrVa 318
 QY 700 -----ACCTCCGCATATAATTCGC 720
 Db 318 lValGlnGluAlaThrLeuProIleIleLeuGln 329
 RESULT 11
 T43213
 ENBP1 protein - barrel medic
 N/Alternate names: chloroplast-localised DNA-binding protein PD3 homolog
 C/Species: Medicago truncatula (barrel medic)
 C/Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
 C/Accession: T43213
 R/Christiansen, H.; Barker, D.G.
 submitted to the EMBL Data Library, October 1997
 A/Description: ENBP1 is homologous to a chloroplast-localised protein from pea, PD3.
 A/Reference number: Z22340
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-1701 <CHR>
 A/Cross-references: UNIPROT:082022; EMBL:AJ002479; NID:el1317985; PID:el1317986; PIDN:CAAC
 A/Experimental source: cultivar Jemalong
 C/Genetics:
 A/Gene: ENBP1
 A/Introns: 856/3; 900/2; 933/3; 978/1; 1000/2; 1220/3; 1329/3; 1585/3; 1610/1; 1638/3; 1
 Alignment Scores:
 Pred. No.: 0.122 Length: 1701
 Score: 115.50 Matches: 70
 Percent Similarity: 32.77% Conservative: 47
 Best Local Similarity: 19.61% Mismatches: 129
 Query Match: 5.89% Indels: 111
 DB: 2 Gaps: 13
 US-10-046-935-2234 (1-1116) x T43213 (1-1701)
 QY 244 AGGGAGCAACCAACCAAAAGCAGAGTCCCGCCAGCCCTCAGAGAATTCGTGACTGAT 303
 Db 628 LysGlySerLysAsnLysLys----- 634
 QY 304 TCCAACTCCGATTCAGAGATGAAGATGAATTTTGGAGAAAAG----- 354
 Db 635 LysAsnIleAlaGlyGluAspGlyAsnLysLeuHisLysGluLysLysArgGlyTrp 654
 QY 355 -----GCTTAAATATAAGCAAAACAAAGCAATCTTCGAAATCATGTCT 402
 Db 655 ProLysGlyPheCysLeuLysProLysGluIleAlaAlaArgLeuAspGluLysIleGlu 674
 QY 403 GAATTAGAAAGCTTCCTCGCTCGTTCGTGGAAGACATCCCTCCAGGCTCCGACTCA 462
 Db 675 ArgArgGlyArgProLysGlySer-----GlyMetLysProLysGluThrAlaValGln 692
 QY 463 CAATCAGGAGACCGGAGCGGTACATTCGCGGTGTTGCTCCAGGAGA----- 513
 Db 693 LeuAspAlaLysIleGluArgGlyArgProLysGlyAlaGlyLysLysProLysGlu 712
 QY 514 -----AACCTGAACCGAGAGCTCGTCTCTT----- 540
 Db 713 IleValValArgLeuAspThrLysIleGluArgArgGlyArgProLysGlySerGlyLys 732
 QY 541 -----ACCAAGTCAAGGTCCTCGG 558
 Db 733 LysGlnLysGluValAlaSerGlnLeuAlaLeuGlnIleGluSerGlnLysSerThrArg 752

QY 559 ATCTCGGGTCCCTTCGACGCTCTA----- 582
 Db 753 ValAspGlyAlaLeuSerThrIleValProHisLysHisIleGlnGluSerIleSer 772
 QY 583 -----CCCATGGAGAGCCAGTTTGTGARKMTWMTACATGTTGGTGAAGAG 630
 Db 773 ProLeuLysAspProValAsnLysGluGluLysSerAspPheValLeuGluCysSerLys 792
 QY 631 -----AGGAAGACCGTGGATGCTACATCAATGAAGATGACCTGCCAGAGC 678
 Db 793 AspSerGlyIleGluLysIleThrLysGlyLeuMetSerLysSerGlyAspValHisLys 812
 QY 679 CGTCGCTCCAGATCATCCGTCACCTTCGCGCATATAAATTCGCCACAGTGAAGAAATAC 738
 Db 813 ArgCysSerGluArgLeuArgThrLeuLeuThrAspHisLysAsnSerGlnAspValGlu 832
 QY 739 GAGGAGAGTTGGAGAACGCTCTGCAGCAATTCGAGAGAAAGATATATACCGTTCACGTG 798
 Db 833 ValGlu-----GluThrPheCysGluAsnGluValGluGluAlaIleAspHisGluLeu 850
 QY 799 GGCTCTACT----- 810
 Db 851 GluSerSerAspLeuMetGlyGluProGluThrLysLysGluProArgAsnLeuArgCys 870
 QY 811 CATCAATGCCGTGAGAGACT-----ATTGATACCAAAACAAACAAACGTAACCCCA 861
 Db 871 HisGlnCysTrpLysLysSerArgThrGlyIleValValCysThrLysCysLys----- 888
 QY 862 GACTGCTGGGGCTTCGAGCCAGTTCGTGCCCCCTGCTTCGAAACCGTTATGTTGAA 921
 Db 889 -----ArgLysLysTrpCysTrpGluCysIleAlaLysTrpTrpGlnAsp 903
 QY 922 GAGGTCAGGAGTCTCTGCTGATCCGAACCTGGCATTCGCCCTTCGAGGAATCTGC 981
 Db 904 LysThrArgGluGluLe-----GluThrAlaCysProPheCysLeuAspTyrCys 920
 QY 982 AACTGCAGTTTCTGCCGCGCAGCAGATGGACGGTGTGCGACTGGGGTCCCTGTGTATT 1041
 Db 921 AsnCysArgLeuCysLeuLysLysThr----- 930
 QY 1042 GCCAAATATCATGCTTTGGGAATGTGCATGCTACTTCAAAGCCTGAAA 1092
 Db 931 SerThrMetAsnGlyAsnGlyGluAlaAspValLysLeuGlnLys 947
 RESULT 12
 G59436
 KIAA1304 protein [imported] - human
 C/Species: Homo sapiens (man)
 C/Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 09-Jul-2004
 C/Accession: G59436; H59436
 R/Nagase, T.; Kikuno, R.; Ishikawa, K.I.; Hirose, M.; Ohara, O.
 DNA Res. 7, 65-73, 2000
 A/Title: Prediction of the coding sequences of unidentified human genes. XVI. The complete
 A/Reference number: G59436
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-1051 <NAG>
 A/Cross-references: UNIPROT:Q7Z6B7; GB:BAA92542; PID:g7242963; PIDN:BAA92542.1
 R/Ohara, O.; Nagase, T.; Kikuno, R.
 submitted to GenBank, January 2000
 A/Reference number: H59436
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-1051 <OHA>
 A/Cross-references: GB:BAA92542; PID:g7242963; PIDN:BAA92542.1
 Alignment Scores:
 Pred. No.: 0.127 Length: 1051
 Score: 115.00 Matches: 77
 Percent Similarity: 31.70% Conservative: 46
 Best Local Similarity: 19.85% Mismatches: 129

Db 318 LysLysAlaGlnGluAsnAspValLysLysGlnLeuThrLysGlnLysSerMetProAla 337
QY 565 GGG-----TCCCTTGAGCTCTACCCATCGGAGAGCCAGT 600
Db 338 GlyGluArgTyrSerGlnGluSerSerGlyLeuAspArgProLeuLysGlnLys 357
QY 601 TTKGARKMTMMGTACATGTTG-----GTGAGAAAGAGAGACC 639
Db 358 LeuAspGlySerValThrValArgAspGlyTyrAspThrThrIleLeuGlnAsnIleThr 377
QY 640 GTGGATGCTACATGAATGAAGATGACCTGCCAGAGCCGCTCCAGATCATCCGTTG 699
Db 378 SerAspGly-----LysLysAspThrGlySerPheLysArgProArgAspLysVal 394
QY 700 ACCCTTCCGCATATAATTCCGCCAGTGAAGAAATACAGAGAGAGAGTGGAGAAGTC 759
Db 395 Thr-----TleGluGluValProGluLysArgSerPheVal 407
QY 760 TGCAGAAATTCGAGAGAGATATATAACCTTCACTGGCTCTACTTGTTCATCAATGC 819
Db 408 -----LysAsnArgAspLeuValSerValSerGluGlyLysThrThrLysThrVal 425
QY 820 CGTCAGAGACTATTGATACCAACAAACAC-----849
Db 426 ThrGluLysGlyLysSerLysLysProSerPheGlyArgAlaGluAspLysMetSerAla 445
QY 850 -----TGCAAGAAACCCAGACTGCTGGGC 873
Db 446 AspAspAsnGluArgAsnTyrGlnValThrGluValCysArgArgProAspAlaArgVal 465
QY 874 GTT-----CGAGGCCAGTTCTGTGCCCCCTGCTTCGAAACCGTTATGTT 918
Db 466 ValLysSerLeuAspAlaArgMetTrpPheCysLysAsnThrIle-----480
QY 919 GAAGAGTCAGGATGCTCTGCTG-----GATCCGAAGCTGCATTCGCCGCTTGTGCA 972
Db 481 -----LeuLysGlyAlaValLeuPheTyrSerProSer-----HisCysMet-----494
QY 973 GGAATCTGCAACTGCAGTTCTTCTGCCGCGCAGAGATGGACGGTGTGCGACTGGGGTCTT 1032
Db 495 ---LeuTyrSerCysSer-----AspValLeu 502
QY 1033 GTGTATTAGCCAAATATCAT-----GGCTTTGGAAATGTGCATCGCTACTTGA 1083
Db 503 SerPheSerSerTyrTyrValThrThrMetLeuGlyAsnIleGluProTrpGluGlu 522
QY 1084 AGCCTGMAACAG 1095
Db 523 SerMetArgGlu 526

RESULT 14
T10955
early nodulin binding protein 1 - spring vetch
C:Species: Vicia sativa (spring vetch, tare)
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: T10955
R:Christiansen, A.; Hansen, A.C.; Vijn, I.; Pallisgaard, N.; Larsen, K.; Yang, W.C.; Biss
submitted to the EMBL Data Library, December 1995
A:Description: A novel type of DNA binding protein interacts with a conserved sequence i
A:Reference number: Z17228
A:Accession: T10955
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1641 <CHR>
A:Cross-references: UNIPROT:Q41700; EMBL:X95995; NID:gl360633; PID:e225826
C:Genetics:
A:Note: ENBP1
C:Keywords: DNA binding

Alignment Scores:
Pred. No.: 0.199 Length: 1641
Score: 113.00 Matches: 73
Percent Similarity: 32.50% Conservative: 44

Best Local Similarity: 20.28% Mismatches: 125
Query Match: 5.76% Indels: 118
DB: 2 Gaps: 17
US-10-046-935-2234 (1-1116) x T10955 (1-1641)
QY 67 GTGAAGTTGATTTCCATGGAACCTCGTCATCTCTGATGACAGTTGTGACAGCTTTGCT 126
Db 585 IleLysLeuIleGlyMetGluSer-----GluAlaValSerValGly 600
QY 127 TCTGATATTTTGAACACAGAGGCTGCAGTCAGTTCCGGAAGGCTGTAGGACCCGACG 186
Db 601 AspGlnGluLeuValValGlnProLeuValLysValArg-----613
QY 187 CAGTGCAGCACTCTGGACCTCTCAGGTGGCGATGAAGTTTCCAGCGGAGTACCAG 246
Db 614 -----PheArgMetLeuAsnProLysMetGlyArgPro-----Lys 625
QY 247 GGAGCAACCAACAAAGACAGAGTCCGCCACCCCTCAGAGATTTCTGTGATGATCC 306
Db 626 GlySerLysAsnLys-----Lys 632
QY 307 AACTCCGATTCAAGAGATGAAGTGAATGAATTTTGGAGAAAG-----354
Db 633 AsnValAspGlyGluAlaGluAsnGlyLeuHisLysGluGlyLysArgGlyArgPro 652
QY 355 -----GCTTTAAATATAAGCAA-----AACAAAGCAATGCTTCAAACTCATG 399
Db 653 LysGlySerValAsnAsnProLysGluThrGlyAsnGluLysIleAlaLysGlyLeuVal 672
QY 400 TCTGAATTAGAAAGCTTCCCTGGCTGCTTCCGTGGAAGACATCCCTCCAGGCTCCGAC 459
Db 673 SerGlu-----Ser 675
QY 460 TCACATCAAGAGACCGGAGAGCGTACATTCCTCCGGGTGTGCTTCCAGGAGAACCT 519
Db 676 SerAsnValHisLysIleGluArgArgGlyArgPro-----LysGlySerAlaProAsnPro 694
QY 520 GAACGGAGAGCTCGTCTCTT-----ACCAGGTCAAGGTCCCGGATCCCTC 564
Db 695 LysGluAsnAlaSerArgLysAspAlaGluIleGluArgGluLysAsnThrHisValTyr 714
QY 565 GGGTCCCTTGAC---GCTCTACCATGGAGAGCCAGTTTKGARKMTMMGTACTGTTG 621
Db 715 GlyIleLeuSerThrThrMetProHisLysHisIleHisGluGluSerIleLeuLeu 734
QY 622 -----GTCAGAAAGAGGAGACCGTGGAT-----645
Db 735 GluAspHisValAsnLysLysAspAlaAspPheValLeuGluCysSerLysGluSer 754
QY 646 -----GGCTACATGAATGAAGATGACCTGCCAGAGACCGT---681
Db 755 GlyIleGluLysIleAlaLysGlyLeuValSerGluSerAspAsnValHisLysThrHis 774
QY 682 -----CGCTCCAGATCATCGTACCGTCCCGTCCGATATA 714
Db 775 AspValGluValGlyAspIlePheTyrGluLysGluValLysGluThrIleAspHisArg 794
QY 715 ATCCGCCCGAGTGAAGAAATACAGAGGAGGTTGGAGACGCTCTGAGCAATTCGA 774
Db 795 LeuGluProSerAspMetLeuGlyAspCysGluThrLysLysGluProArgAsnSerArg 814
QY 775 GAGAAGATATATAACCGTTCACTGGGCTCTACTTGTTCATCAATCGCTCCAGAGACT---831
Db 815 -----CysHisGlnCysTrpLysLysSerArg 823
QY 832 -----ATTGATACCAAAACAACTGAGAAACCCAGACTGCTGGGGCGTTCGAGCCAG 885
Db 824 ThrGlyLeuValValCysSerLysCys-----LysLysLys 836
QY 886 TTCTGTGGCCCTGCTTCGAAACCGTTATGTTAGAGAGTCCAGGATGCTCTGCTGGAT 945
Db 837 TyrCysTyrGluCysValAlaLysTrpTyrGlnAspLysThrArgGluGluIle-----854

QY 946 CGGAAGTGGCATTGCCCGCTTGTGCGAGGATCTGCAACTGCAGTTTCTGCGCGGACGCA 1005
Db 855 ---GluThrAlaCysProPheCysLeuAspTyrCysAsnCysArgMetCysLeuLyslys 873
RESULT 15
A47392
Chromodomain-helicase-DNA-binding protein, CHD-1 - mouse
X:Alternate names: KYP protein
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Dec-2002
C:Accession: A47392; S21568
R:Delmas, V.; Stokes, D.G.; Perry, R.P.
Proc. Natl. Acad. Sci. U.S.A. 90, 2414-2418, 1993
A:Title: A mammalian DNA-binding protein that contains a chromodomain and an SNF2/SWI2-1
A:Reference number: A47392; MUID:93211972; PMID:8460153
A:Accession: A47392
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1711
A:Experimental source: S194 plasmacytoma cells
A:Note: sequence inconsistent with the nucleotide translation
A:Note: sequence extracted from NCBI backbone (NCBIN:128272, NCBI:P.128273)
R:Delmas, V.; Perry, R.P.
submitted to the EMBL Data Library, May 1992
A:Description: KYPB, a mammalian protein that contains the SNF2/SWI2 helicase domain also
A:Reference number: S21568
A:Accession: S21568
A:Molecule type: mRNA
A:Residues: 772-1711 <DE2>
A:Cross-references: EMBL:X66028
C:Superfamily: chromodomain helicase CHD1; chromobox homology
C:Keywords: DNA binding
F:293-336/Domain: chromobox homology <CB1>
F:387-427/Domain: chromobox homology <CB2>

Alignment Scores:
Pred. No.: 0.27 Length: 1711
Score: 111.50 Matches: 82
Percent Similarity: 38.04% Conservative: 58
Best Local Similarity: 22.28% Indels: 95
Query Match: 5.68% Gaps: 18
DB: 1

US-10-046-935-2234 (1-1116) x A47392 (1-1711)

QY 79 TCATGGAACCTCGTCATCTCTGATGACAGTGTGACAGCTTGTCTCTGATATTTT 138
Db 39 SerSerAspGlySerSerSerGlnSerGlySerSerAsp-----SerAspSerGly 55
QY 139 GCAACACAGAGCTGCAGTCAGTTCGGGAGGCTGTAGGACCGCAGCCAGTCGAGGCAC 198
Db 56 SerAspSerGlySerGlnSerGluSerGluSerAspThrSerArgGluAsnLysValGln 75
QY 199 TCTGGACCTCTCAGGCTG---GCGATGAAGTTCACGCGCGAGT-----240
Db 76 AlaLysProProLysValAspGlyAlaGluPheTrpLysSerSerProSerIleLeuAla 95
QY 241 ---ACCAGGGAGCAACCAACAAAAA-----GCAGAGTCGCCAGCCGCCC---282
Db 96 ValGlnArgSerAlaMetLeuArgLysGlnProGlnGlnAlaGlnGlnArgProAla 115
QY 283 TCAGAGAAATCTGTGACTGATTCCACTCCGATTCGATGAGATCAAGATGAATTTT 342
Db 116 SerSerAsnSerGlySerGluGluAspSerSerSerSerGluAspSer-----131
QY 343 TTGAGAAAAAGGCTTTAAATATATAAAGCAAAAGCAATGCTTCGAAAACTCATGCT 402
Db 132 ---AspAspSerSerSerGlyAlaLysArgLysLys-----His 143
QY 403 GAATTAGAAAGCTTCCTGGCTGCTTCGTTGGTGGAGAGATCCCTCCAGGCTCCGACTCA 462
Db 144 AsnAspGluAspTrpGlnMetSerGlySerGlySerProSerGlnLeuGlySerAspSer 163

QY 463 CAATCAAGGAGACCGCGA---AGCGGTACATTCCCGGTGTGTTGCTTCCAGGAGAAACCCCT 519
Db 164 GluSerGluGluGluArgAspLysSerSerCysAspGlyThrGluSerAspTyrGluPro 183
QY 520 GAACGAGAGAGCTCGT-----CCTCTTACAGGTCAAGTCC-----CGG 558
Db 184 LysAsnLysValArgSerArgLysProGlnAsnArgSerLysSerLysAsnGlyLysLys 203
QY 559 ATCTCTCGGTCCTTGCAGCTCTACCCATGGGAGAGCCAGTTTGARKMTWGTACATG 618
Db 204 IleLeuGly-----206
QY 619 TTGGTGAAGAAAGAGAGACCGTGTGATGGCTACATGAATGAAGATGACCTG-----669
Db 207 -----GlnLysLysArgGlnLysSerSerSerGluAspGluAspGluAspLysTrpAsp 224
QY 670 -----CCCAGAACCGTCGCTCCAGATCATCCGTGACCCCTTCCGCATATAATTCGCCCA 723
Db 225 AsnAspLysArgSerSerArgArgGlnAlaThrValAsnValSerTyrLysGluAspGlu 244
QY 724 GTGGAAGAAATTCACAGAGAGAGTGGAGAACGTTGAGCAACTCTCGAGAGAAGATA 783
Db 245 GluMetLysThrAspSerAspLeuLeuValCysGlyGluAspValProGlnPro 264
QY 784 TATAACCGTTCACCTGGCTCTACTTGTTCATCAATCCGTCGAGAGACTATTGATACCAA 843
Db 265 GluAspGluGluPhe-----GluThrIleGluArgVal 275
QY 844 ACAAACTGCAGAAACCCAGACTGTGGGGCGTTCGAGGCCAGTCTGTGGCCCTGCTT 903
Db 276 MetAspCysArg-----ValGlyArgLysGlyAlaThrGlyAlaThr 289
QY 904 CGAAACCGTGTGTAAGAGTCAAGGATGCTCTGTGGATCCGAACTGGCATTCGCCG 963
Db 290 ThrThrIleTyrAlaValGluAlaAspGly-----AspPro-----301
QY 964 CCTTGTGCGAGGAATCTGCAACTGCAGTTTCTGCCGCGAGAGATGGACGGTGTGCACT 1023
Db 302 -----AsnAlaGlyPheGluArgAsnLysGlu-----Pro 311
QY 1024 GGGGTCTTGTGTATTTAGCCAAATATCATGCTTTGGGAATGTGCAT-----GCC 1074
Db 312 GlyAspIleGlnTyrLeuIleLysTrpLysGlyTyrSerHisIleHisAsnThrTrpGlu 331
QY 1075 TACTTGAAGAGCTGAACAGGAA 1098
Db 332 ThrGluGluThrLeuLysGln 339

Search completed: October 12, 2004, 15:19:50

Job time : 43 secs

Blank

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 12, 2004, 15:05:41 ; Search time 171.5 Seconds
(without alignments)
7488.258 Million cell updates/sec

Title: US-10-046-935-2234

Perfect score: 1962

Sequence: 1 atggacgctgcgcggtgcc.....aattgaaatgaacataa 1116

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 3650362

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-O=/cn2_1/USPTO_spool_p/US10046935/runat_12102004_155225_15243/app_query.fasta_1.1287
-DB=Uniprot_02 -QFMT=fastan -SUFFIX=rxp -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10046935 @CN2_1_1.183 @runat_12102004_155225_15243 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Uniprot 02:.*
1: uniprot_sprot:.*
2: uniprot_trembl:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1929	98.3	371	Q9BWT1	Q9bwt1 homo sapien
2	1879.5	95.8	450	Q9GSP5	Q9gsp5 homo sapien
3	1612.5	82.2	382	Q9D0M2	Q9d0m2 m mus muscu
4	1609.5	82.0	382	Q6NZE5	Q6nze5 mus muscu
5	1609.5	82.0	382	Aah66169	Aah66169 mus muscu
6	1607.5	81.9	382	Q8C1A0	Q8c1a0 mus muscu
7	1155	58.9	220	Q96BV8	Q96bv8 homo sapien
8	786	40.1	207	Q8NBY9	Q8nby9 homo sapien
9	752	38.3	442	Q96C70	Q96c70 homo sapien
10	752	38.3	453	Q86YT0	Q86yt0 homo sapien
11	752	38.3	453	Aah25242	Aah25242 homo sapi
12	752	38.3	454	Q96GN5	Q96gn5 homo sapien
13	743	37.9	453	Q6P1L4	Q6p1l4 homo sapien
14	743	37.9	453	Aah32576	Aah32576 homo sapi
15	736	37.5	438	Q922M5	Q922m5 mus muscu
16	602.5	30.7	202	Q9NPV2	Q9npv2 homo sapien

17	289.5	14.8	462	2	Q23175	Q23175 arabidopsis
18	250.5	12.8	515	2	Q9FXE1	Q9fxe1 arabidopsis
19	233.5	11.9	544	2	Q9FEV3	Q9fev3 arabidopsis
20	233.5	11.9	572	2	Q8RW95	Q8rw95 arabidopsis
21	225.5	11.5	495	2	Q9FYG6	Q9fyg6 arabidopsis
22	216.5	11.0	297	2	Q9AURO	Q9aur0 oryza sativ
23	171.5	8.7	268	2	Q9H9A2	Q9ha2 homo sapien
24	170	8.7	555	2	Q80469	Q80469 arabidopsis
25	146	7.4	851	2	Q6S384	Q6s384 arabidopsis
26	142	7.2	950	2	Q04024	Q04024 arabidopsis
27	132.5	6.8	566	2	Q6Q150	Q6q150 arabidopsis
28	132.5	6.8	566	2	AAS75310	Aas75310 arabidops
29	132.5	6.8	570	2	Q9LV75	Q9lv75 arabidopsis
30	131.5	6.7	906	2	Q48794	Q48794 arabidopsis
31	129	6.6	754	1	PP1G_HUMAN	PP1G_HUMAN
32	124.5	6.3	1007	1	PR4B_MOUSE	PR4b_mouse muscu
33	124.5	6.3	1007	2	Q8BND8	Q8bnd8 mus muscu
34	123.5	6.3	1458	2	Q9HE72	Q9he72 neurospora
35	122	6.2	820	2	Q60585	Q60585 homo sapien
36	121.5	6.2	497	2	Q8CSG1	Q8csg1 mus muscu
37	119.5	6.1	1150	2	Q9W0H4	Q9w0h4 drosophila
38	118.5	6.0	822	2	Q6DFC7	Q6dfc7 xenopus lae
39	118.5	6.0	879	2	Q80Y24	Q80y24 mus muscu
40	118	6.0	2759	2	Q9VID9	Q9vid9 drosophila
41	117.5	6.0	595	2	Q9ZRZ9	Q9zrz9 arabidopsis
42	117.5	6.0	827	2	Q9FNM7	Q9fnm7 arabidopsis
43	117	6.0	1001	2	Q75DD5	Q75dd5 ashbya goss
44	117	6.0	1001	2	AAS0862	Aas0862 ashbya go
45	116.5	5.9	514	2	Q9NWD4	Q9nwd4 homo sapien

ALIGNMENTS

RESULT 1

Q9BWT1 ID Q9BWT1 PRELIMINARY; PRT; 371 AA.
AC Q9BWT1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE C-Myc target JPO1 (Cell division cycle associated protein 7, isoform
DE 2) (Hypothetical protein FLJ14736).
GN Name=JPO1; Synonyms=CDCA7;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21611205; PubMed=11598121;
RA Prescott J.E., Osthus R.C., Lee L.A., Lewis B.C., Shim H.,
RA Barrett J.F., Guo Q., Hawkins A.L., Griffin C.A., Dang C.V.;
RT "A novel c-Myc-responsive gene, JPO1, participates in neoplastic
RT transformation.";
RL J. Biol. Chem. 276:48276-48284 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Hsien F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISU8Pancreas;
 RA Strausberg R.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 EX PubMed-14702039;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto K., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,
 RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
 RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
 RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yanazaki M.,
 RA Nimomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
 RA Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y., Ishida S.,
 RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
 RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
 RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
 RA Mubashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
 RA Yoshioka K., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
 RA Moriya S., Momiya H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Ohtani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs";
 RL Nat. Genet. 36:40-45 (2004).
 DR EMBL; AY029179; RAK31591.1; -;
 DR EMBL; BC027966; AAH27966.1; -;
 DR EMBL; AK027642; BAB55258.1; -;
 KW Cell division.
 SQ SEQUENCE 371 AA; 42573 MW; 30A244E3057D9C43 CRC64;

Alignment Scores:

Pred. No.: 4,6e-153 Length: 371
 Score: 1929.00 Matches: 364
 Percent Similarity: 98.11% Conservative: 0
 Best Local Similarity: 98.11% Mismatches: 7
 Query Match: 98.32% Indels: 0
 DB: 2 Gaps: 0

US-10-046-935-2234 (1-1116) x Q9BWT1 (1-371)

QY 1 ATGGAGCGCTCGCGCGTCCGCGGAGAGATCTCAGAGTAAAGAACTTAAAGAAATTC 60
 DB 1 MetAspAlaArgArgValProGlnLysAspLeuArgValLysAsnLeuLysPhe 20
 QY 61 AGATATGTGAAGTTGATTTCCATGGAAACCTCGTATCTCTGATGACAGTTGTGACAC 120
 DB 21 ArgTyrValLysLeuIleSerMetGluThrSerSerSerSerSerSerSerSerSer 40
 QY 121 TTTCCTCTTGATATTTTGCACACAGAGCTCAGTTCGGGAGCGCTGTAGAC 180
 DB 41 PheAlaSerAspAsnPheAlaAsnThrArgueGlnSerValArgIleGlyCysArgThr 60

QY 181 CGCAGCCAGTGCAGGCACCTCTGGACCTCTCAGGGTGGCGATGAAGTTTCCAGCGCGGAGT 240
 DB 61 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 80
 QY 241 ACCAGGGGAGCAACCAACAAAAAGCAGAGTCCCGCAGCCCTCAGAAATTCGTGACT 300
 DB 81 ThrArgGlyAlaThrAsnLysAlaGluSerArgGlnProSerGluAsnSerValThr 100
 QY 301 GATTCCAACTCCGATTCCAGAGATGAAGTGAATGAATTTTTCGAGAAAGGCGCTT 360
 DB 101 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 120
 QY 361 AATATAAGCAAAACAAACCAATCTCTGCAAACTCATGTCTGAATTAAGAAGCTTCCT 420
 DB 121 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140
 QY 421 GGCTCGTTCCGTGGAGACATCCCTCCAGGCTCCGACTCACATCATCAAGAGACGGCA 480
 DB 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 160
 QY 481 AGGCCTCATCTCCCGGCTGTTGCTTCCAGGAGAAACCCCTGAACGAGAGCTCGTCTCT 540
 DB 161 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgArgAlaArgProLeu 180
 QY 541 ACCAGTCAAGTCCCGGATCTCGGTCTCCCTTGAAGCTCTACCATCGRGAGAGCCAGT 600
 DB 181 ThrArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMetGluGluGlu 200
 QY 601 TTGARKMTMWTACATCTTGGTGAAGAGGAGAGACCGTGATGGCTACATGAATGAA 660
 DB 201 GluGluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 220
 QY 661 GATGACCTGCCAGAGCGCTCGCTCCAGATCATCCGTGACCTTCCGATATTAATTCG 720
 DB 221 AspAspLeuProArgSerArgArgSerArgSerValThrLeuProHisIleArg 240
 QY 721 CCAGTGAAGAAATTCAGAGAGAGAGTGGAGAACGCTCGAGCAATTCGAGAGAG 780
 DB 241 ProValGluGluIleThrGluGluLeuGluAsnValCysSerAsnSerArgGluLys 260
 QY 781 ATATATACCGTTCACTGGCTCTACTTCTCATCAATCCGTCAGAAAGACTATTGATACC 840
 DB 261 IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 280
 QY 841 AAAACAAATCGAGAAACCCAGCTGCTGGGCGCTTCGAGGCGAGTCTGTGGCCCTGC 900
 DB 281 LysThrAsnCysArgAsnProAspCysTrpGlyValArgGlyGlnPheCysGlyProCys 300
 QY 901 CTTGCAAAACCGTTATGTTGTTAGCCAAATATCATGGCTTGGGAATGTGCATGCTTGC 960
 DB 301 LeuArgAsnArgTyrGlyGluGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 320
 QY 961 CGCGCTTTCGAGGAGATCTGCAACTGCAGTTTCTCGCGCAGCAGAGATCGAGTGTGCG 1020
 DB 321 ProProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 340
 QY 1021 ACTGGGCTCCTTGTGTATTAGCCAAATATCATGGCTTGGGAATGTGCATGCTTGC 1080
 DB 341 ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 360
 QY 1081 AAAAGCTGAAACAGAAATTTGAAATGCAAGCA 1113
 DB 361 LysSerLeuLysGlnGluPheGluMetGlnAla 371
 RESULT 2
 Q96SP5 PRELIMINARY; PRT; 450 AA.
 ID Q96SP5;
 AC Q96SP5;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Hypothetical protein FLJ14722.
 OS Homo sapiens (human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

RX PubMed=1470239;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,
 RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
 RA Sudo H., Hosoliri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
 RA Abe K., Kamihara K., Katsura N., Sato K., Tanikawa M., Yamazaki M.,
 RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
 RA Tanai H., Kimata M., Watanabe M., Hirao K., Chiba Y., Ishida S.,
 RA Ono Y., Takiguchi S., Watanabe S., Yoshida M., Hotuta T., Kusano J.,
 RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
 RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
 RA Mutsaers K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki M., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Satoh T., Kikuchi H., Masuho Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isegai T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs".
 RL Nat. Genet. 36:40-45 (2004).
 DR EMBL; AK027628; BAB5245.1; -;
 DR Genbank; HGNC:14628; CDCA7.
 SQ SEQUENCE 450 AA; 51417 MW; D86A79E7981B773D CRC64;

Alignment Scores:

Pred. No.: 6,93e-149 Length: 450
 Score: 1879.50 Matches: 364
 Percent Similarity: 80.8% Conservative: 0
 Best Local Similarity: 80.8% Mismatches: 7
 Query Match: 95.8% Indels: 79
 DB: 2 Gaps: 1

US-10-046-935-2234 (1-1116) x Q96SP5 (1-450)

QY 1 ATGGAGCTCGCGGTCGCGGAGAGAGTCTCAGAGTAAGAGACTTAAGAAATTC 60
 Db 1 MetAspAlaArgValProGlnLysAspLeuArgValLysLysAsnLeuLysPhe 20
 QY 61 AGATATGTGAAGTTGATTTCCATGGAAACCTCGTCATCTCTGATCAGAGTTGTGACAGC 120
 Db 21 ArgTyrValLysLeuLeuSerMetGluThrSerSerSerSerSerSerSerSerSerSer 40
 QY 121 TTTCCTCTCTGATAAATTTTGCACACG----- 147
 Db 41 PheAlaSerAspAsnPheAlaThrLysProLysPheArgSerAspIleSerGluGlu 60
 QY 147 ----- 147
 Db 61 LeuAlaSerValPheTyrGluAspSerAspAsnGluSerPheCysGlyPheSerGluSer 80
 QY 147 ----- 147
 Db 81 GluValGlnAspValLeuAspHisCysGlyPheLeuGlnLysProArgProAspValThr 100
 QY 147 ----- 147

Db 101 AsnGluLeuAlaGlyIlePheHisAlaAspSerAspGluSerPheCysGlyPheSer 120
 QY 148 -----AGGCTGCAGTCAGTTCGGGAAGGCTGTAGGACCCGC 183
 Db 121 GluSerGluIleGlnAspGlyMetArgLysGlnSerValArgGluGlyCysArgThrArg 140
 QY 184 AGCCAGTGCAGGCACCTCTGAGGCTGGCGATGAAGTTTCCAGCGCGAGTAGTACC 243
 Db 141 SerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSerThr 160
 QY 244 AGGGAGCAACCAACAAACAAAGACAGAGTCCCGCCAGCCCTCAGAGAAATCTGTGACTGAT 303
 Db 161 ArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThrAsp 180
 QY 304 TCCAACTCCGATTCAGAGATGCAAGTGAATTTTTCGAGAAAGGCGCTTTAAAT 363
 Db 181 SerAsnSerAspSerGluAspLysSerGlyMetAsnPheLeuGluLysArgAlaLeuAsn 200
 QY 364 ATAAAGCAAAACAAACGAATGCTTCGAAAACCTCATGTCTGAATTTAGAAGCTTCCTCGC 423
 Db 201 IleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPheProGly 220
 QY 424 TCGTTCCTGGAGAGACATCCCTCCAGGCTCCGACTCACATCAAGAGACCGGAGAG 483
 Db 221 SerPheArgGlyArgHisProLeuProLysSerAspSerGlnSerArgProArgArg 240
 QY 484 CGTACATTCCTCGGATCTCGGGTCCCTTGACGCTCTACCCATCGRGAGAGCCAGTTTK 543
 Db 241 ArgThrPheProGlyValAlaSerArgArgAsnProGluArgAlaArgProLeuThr 260
 QY 544 AGGTCAAGGTCCTCGGATCTCGGGTCCCTTGACGCTCTACCCATCGRGAGAGCCAGTTTK 603
 Db 261 ArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMetGluGluGluGlu 280
 QY 604 GARKMTMGTTACATOTTGGTGAAGAGAGAGACCGTGGATGGCTACATGATGAAGT 663
 Db 281 GluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGluAsp 300
 QY 664 GACCTGCCAGAGAGCGTCCGCTCCAGATCATCCGTCGACCTTCCGATATAATTCGCCCA 723
 Db 301 AspLeuProArgSerArgArgSerArgSerValThrLeuProHisIleIleArgPro 320
 QY 724 GTGGAAGAAATTCAGAGAGAGGAGTTGGAGAACCGTCTGCAGCAATTCACGAGAGAGATA 783
 Db 321 ValGluGluIleThrGluGluGluLeuGluAsnValCysSerAsnSerArgGluLysIle 340
 QY 784 TATAACCGTTCACTGGGCTCTACTTGTCTCATCAATGCCGTCAGAGACTTATTCATACCAA 843
 Db 341 TyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThrLys 360
 QY 844 ACAAACTGCAGAAACCCAGACTGCTGGGCGGTTCCGAGGCGAGTCTGTGGCCCTGCCCTT 903
 Db 361 ThrAsnCysArgAsnProAspCysTyrGlyValArgGlyGlnPheCysGlyProCysLeu 380
 QY 904 CGAAACCGTTATGGTGAAGAGGTCAGGGATGCTCTCTGGATCCGAACTGGCATTCGCCCG 963
 Db 381 ArgAsnArgTyrGlyGluValArgAspAlaLeuLeuAspProAsnTrpHisCysPro 400
 QY 964 CTTGTGAGGAATCTGCAACTGCACTGAGTTCTCCCGGACAGAGATGGAGGTTGGGACT 1023
 Db 401 ProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAlaThr 420
 QY 1024 GGGGCTCTGTGCTATTAGCCAAATATCATGGCTTTGGGAATGTGCATCCCTACTTGA 1083
 Db 421 GlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeuLys 440
 QY 1084 AGCTGAAACAGGAATTTGAAATGCAAGCA 1113
 Db 441 SerLeuLysGlnGluPheGluMetGlnAla 450

RESULT 3
 Q9DOM2
 ID Q9DOM2 PRELIMINARY; PRT; 382 AA.

Q9D0M2;
 01-JUN-2001 (TrEMBLrel. 17, Created)
 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Mus musculus 10 days embryo whole body cDNA, RIKEN full-length
 DE enriched library, clone:2610002110 product:similar to C-MYC TARGET
 DE JPO1 (CDNA FJ14736 FIS, CLONE NT2RP3002181) (Mus musculus adult male
 DE tongue cDNA, RIKEN full-length enriched library, clone:2310021G01
 DE product:similar to C-MYC TARGET JPO1).
 GN Name=Cdc47;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Tongue, and Whole body;
 RX MEDLINE=99279253; PubMed=10349636;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Tongue, and Whole body;
 RX MEDLINE=21085660; PubMed=11217851;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Tongue, and Whole body;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Tongue, and Whole body;
 RX MEDLINE=20499374; PubMed=11042159;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Tongue, and Whole body;
 RX MEDLINE=20530913; PubMed=11076861;
 RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RX MEDLINE=99279253; PubMed=10349636;
 RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
 RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
 RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
 RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
 RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
 RA Muramatsu M., Hayashizaki Y.;

RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 [7]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Tongue;
 RA Adachi J., Aizawa K., Akahira S., Akimura T., Aono H., Arai A.,
 RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
 RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
 RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
 RA Tejima Y., Toya T., Yamamura T., Yamanaka I., Yasunishi A.,
 RA Yoshida K., Yoshino M., Muramatsu M., Hayashizaki Y.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK011289; BAB27519.1; -;
 DR EMBL; AK075882; BAC36027.1; -;
 DR MGD; MGI:1914203; Cdc47.
 SQ SEQUENCE 382 AA; 43837 MW; 51920663588CDD40 CRC64;
 Alignment Scores:
 Pred. No.: 1.76e-126 Length: 382
 Score: 1612.50 Matches: 309
 Percent Similarity: 86.72% Conservative: 24
 Best Local Similarity: 80.47% Mismatches: 36
 Query Match: 82.19% Indels: 15
 Gaps: 2
 US-10-046-935-2234 (1-1116) x Q9D0M2 (1-382)
 QY 1 ATGACGCTCGCGTCCGAGAAAGATCTCAGAGTAAAGAACTTAAAGAAATTC 60
 Db 1 MetGluAlaArgAlaArgGlnLysAlaLeuLysVal---LysAsnLeuLysAspVal 19
 QY 61 AGATATGTGAAGTTGATTTCCATGGAACCTCGTCATCTCTCGATGACAGTTGTGACAGC 120
 Db 20 ArgTyrMetLysLeuLysSerMetGluThrSerSerSerSerSerSerSerSerSerSer 39
 QY 121 TTTCTCTGTGATTAATTTTGCACACACAGGCTGAGTTCGGAGAGGCTGTAGGACC 180
 Db 40 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnLeuAsnArgGluGlyCysArgThr 59
 QY 181 CGCAGCAGTCCAGGCACTCTGGAGCTCTCAGGTTGGCGATGAAGTTTCCAGCGCGAGT 240
 Db 60 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgAsn 79
 QY 241 ACCAGGGAGCAACCAACAAACAAAGAGCAGTCCCGCCAGCCCTCAGAGAAATTCGTGACT 300
 Db 80 ThrArgArgAlaAlaSerLysLysAlaAlaProLysProSerGluSerSerSerSerSer 99
 QY 301 GATTCCCACTCCGATTCA-----GAAGATGAAGTGGATGAT 339
 Db 100 AspSerHisSerAspSerGluGluGluGluGluGluGluGluGluGluGluGluMetAsn 119
 QY 340 TTTTGTGAGAAAGGGCTTTAAATATAAGCAAAACAAAGCAATGCTTGCAGAACTCATG 399
 Db 120 PheLeuGluLysArgAlaLeuAsnLleLysGlnAsnLysAlaMetLeuAlaLysLeuMet 139
 QY 400 TCTGAATTAGAAAGCTTCCCTGGCTCGTTCGTTGGAGAGACATCCCTCCAGGCTCCGAC 459
 Db 140 SerGluLeuGluSerPheProGlyLeuPheSerGlyArgHisSerLeuProGlyHisArg 159
 QY 460 TCACAA---TCAAGGAGACCGGAGGCGTACATTCCTCCGGTGTGTCTCCAGGAGAAC 516
 Db 160 AlaLysAspSerLysSerProArgArgArgThrPheProGlyValAlaIleThrArgArgAsn 179
 QY 517 CCTGAACGAGAGCTCGTCTCTTACAGGTCACAGGTCACAGGTCCTCGGTCCTCTGAC 576
 Db 180 ProGluArgArgThrArgProLeuThrArgSerArgSerArgLleLeuGlySerLeuGly 199
 QY 577 GCTTACCCATG-----GRGAGAGCCAGTTTKGARKMTWGTACATGTTG 621
 |||||||

Db 299 AspCysTyrGlyIleArgGlyGlnPheCysGlyProCysLeuArgAsnArgTyrGlyGlu 318
 QY 922 GAGGTTCAGGATGCTCTGCTGATCCGAATCGATCGCCGCTTGTGTGAGAAATCTGC 981
 Db 319 GluValLysAspAlaLeuLeuAspProAsnTrpHisCysProCysArgGlyIleCys 338
 QY 982 AACTGCAGTTCTTCGCGGAGCAGATGACGCTGTGCGACATGGGCTCTGTGTATTTA 1041
 Db 339 AsnCysSerPheCysArgGlnArgGlyArgCysAlaThrGlyValLeuValTyrLeu 358
 QY 1042 GCCAAATATCATGCTTTCGGATGTCATGCTACTTGAAGGCTTGAACAGCAATTT 1101
 Db 359 AlaArgTyrHisGlyPheGlyAsnValHisAlaTyrLeuLysSerLeuLysGlnGluPhe 378
 QY 1102 GAAATGCCAGCA 1113
 Db 379 GluMetGlnAla 382

RESULT 6

Q8C1A0
 ID Q8C1A0 PRELIMINARY; PRT; 382 AA.
 AC Q8C1A0:
 DT 01-MAR-2003 (TremBLrel. 23, Created)
 DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
 DE Mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched
 DE library, clone:4732431F21 product:similar to C-MYC TARGET JP01 (CDNA
 DE FLJ14736 FIS, CLONE NT2RP3002181).
 GN Name=Cdc47;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Skin;
 RX MEDLINE=99279253; PubMed=10349636;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44 (1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Skin;
 RX MEDLINE=21085660; PubMed=11217851;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690 (2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Skin;
 RX MEDLINE=20499374; PubMed=11042159;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630 (2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Skin;
 RX MEDLINE=20530913; PubMed=11076861;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kusunai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,

RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771 (2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Skin;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
 RA Hori F., Inotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akai H., Takeda Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK028671; BAC26058.1; -;
 DR MGD; MGI:1914203; Cdc47.
 SQ SEQUENCE 382 AA; 43836 MW; 33F004EB500ED9EE CRC64;
 Alignment Scores:
 Pred. No.: 4,64e-126 Length: 382
 Score: 1607.50 Matches: 308
 Percent Similarity: 86.72% Conservative: 25
 Best Local Similarity: 80.21% Mismatches: 36
 Query Match: 81.93% Indels: 15
 DB: 2 Gaps: 5
 US-10-046-935-2234 (1-1116) x Q8C1A0 (1-382)
 QY 1 ATGGACGCTCGCGCTGCGCGAAGATCTCAGAGTAAAGAACTTAAAGAAATTC 60
 Db 1 MetGluAlaArgArgAlaArgGlnLysAlaLeuLysVal--LysAsnLeuLysAspVal 19
 QY 61 AGATATGTGAAGTTGATTTCCATGAAACCTCGTCTATCTCTGTGATGACAGTTGTGACAC 120
 Db 20 ArgTyrMetLysLeuIleSerMetGluThrSerSerSerSerSerSerSerSerSerSer 39
 QY 121 TTTTGCTTCTGATAATTTTGCACACGAGGCTGAGTTCGCGAGGCTGAGGCTGAGGCT 180
 Db 40 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnLeuAsnArgGluGlyCysArgThr 59
 QY 181 CGACCCAGTGCAGCACTCTGGACCTCTCAGGGTGGCGATGAACTTTCCAGCGCGAGT 240
 Db 60 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgAsn 79
 QY 241 ACCAGGGAGGAGCAACCAACAAAGACAGATCCCGCCAGCTTCAGAGAAATTCGTGACT 300
 Db 80 ThrArgArgAlaAlaSerLysLysAlaAlaProProLysProSerSerSerSerSerSer 99
 QY 301 GATTCCAACTCGGATTCA-----GAAGTGAAGTGAATGAT 339
 Db 100 AsnSerHisSerAspSerGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 119
 QY 340 TTTTGGAGAAAAGGCTTTAAATATAAGCAAAACAAAGCAATCTTTCGCAAACTCATG 399
 Db 120 PheLeuGluLysArgAlaLeuAsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMet 139
 QY 400 TCTGATTTAGAAAGCTTCCCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 459
 Db 140 SerGluLeuGluSerPhePhePhePhePhePhePhePhePhePhePhePhePhePhe 159
 QY 460 TCACAA---TCAAGGAGACCGAGGAGGCTACATTCCTCCGGGTGTGCTTCCAGGGAAC 516
 Db 160 AlalysAspSerLysSerProArgArgArgThrPheProGlyValAlaThrArgArgAsn 179
 QY 517 CTTGAACCGAGAGCTCGTCTCTTACCAAGGTCAAGGTCCCGGATCCTCGGTGCCCTTGCAC 576

180 ProGluArgArgThrArgProLeuThrArgSerArgSerArgIleLeuGlySerLeuGly 199
QY 577 GCTTACCCATG-----GRGAGACCGAGTTTGARKMTWGTACATGTTG 621
Db 200 AlaLeuProThrGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 219
QY 622 GTGAGAAAGAGGAGAGCGTGGCTTACATGATGAATGAGTGCCTCCCGAGAGCGGT 681
Db 220 ValArgGlnArgLysSerMetAspSerTyMetAsnAspAspAspValProArgSerArg 239
QY 682 CGCTCCAGATCATCGTGACCGTCCCGCATATAATATCCCGAGTGGAGAGAAATACAGAG 741
Db 240 ArgPro--GlySerMetThrLeuProHisIleArgProValGluGluValThrGlu 258
QY 742 GAGGAGTTGGAGACGCTGCGACCAATTCGAGAGAGATATATAACCGTCTACTGGGC 801
Db 259 GluGluIleArgAsnIleCysSerAsnSerArgGluLysIleTyAsnArgSerLeuGly 278
QY 802 TCTACTTGTCTCAATGCGGTGAGAGAGACTATTGTATACCAAACTGCAGAAACCCA 861
Db 279 SerThrCysHisGlnCysArgGlnLysThrThrAspThrLysThrAsnCysArgAsnPro 298
QY 862 GACTGCTGGGGGTTGAGAGCGAGTCTGTGGCCCGCTCGTAAACCGTTATGCTGAA 921
Db 299 AspCysTrpGlyIleArgGlyGlnPheCysGlyProCysLeuArgAsnArgTyrglyGlu 318
QY 922 GAGGTCAGGAGTCTGCTGTCATCCGAACTGCGCATTCGCCCTTGTGAGGAACTGCG 981
Db 319 GluValLysAspAlaLeuAspProAsnTrpHisCysProCysArgGlyIleCys 338
QY 982 AACTGTCAGTCTTGTGCGGAGAGAGTGTGCGAGTGGGCTCTTGTGTATTATTA 1041
Db 339 AsnCysSerPheCysArgGlnArgAspGlyArgCysAlaThrGlyValLeuValTyLeu 358
QY 1042 GCCAAATATCATGGCTTTGGGAATGTGCATGCTACTTGAAGAGCTGAAACAGAAATT 1101
Db 359 AlaLysTyHisGlyPheGlyAsnValHisAlaTyLeuLysSerLeuLysGlnGluPhe 378
QY 1102 GAAATGCAAGCA 1113
Db 379 GluMetGlnAla 382

RESULT 7

Q96BV8 PRELIMINARY; PRT; 220 AA.
ID Q96BV8; AC Q96BV8;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAUSBERG R.;
RA Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC015124; AAH15124.1; -;
DR InterPro; IPR008972; Cupredoxin.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 220 AA; 25570 MW; C0760EB9AE3899CB CRC64;

Alignment Scores:

Pred. No.: 4,07e-88 Length: 220
Score: 1155.00 Matches: 212
Percent Similarity: 96.80% Conservative: 0
Best Local Similarity: 96.80% Mismatches: 7
Query Match: 58.87% Indels: 0
DB: 2 Gaps: 0

US-10-046-935-2234 (1-1116) x Q96BV8 (1-220)
QY 457 GACTCAGATCAAGAGAGACCGGAGCGGTACATTCCGGGTGTGGTTCACGAGGAAAC 516
Db 2 AspSerGlnSerArgArgProArgArgThrPheProGlyValAlaSerArgArgAsn 21
QY 517 CCTGAACGAGAGAGTCTGCTCTTACACAGGTCAAGGTCCCGATCCTCGGTCCTTCAAC 576
Db 22 ProGluArgArgAlaArgProLeuThrArgSerArgSerArgIleLeuGlySerLeuAsp 41
QY 577 GCTTACCCATGCGAGAGCGCGTTCGARKMTWGTACATGTTGGTTCAGAGAGGAG 636
Db 42 AlaLeuProMetGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 61
QY 637 ACCGTGGATGCTACATGAATGAAGATGACCTGCCAGAACCGTTCGATCATCC 696
Db 62 ThrValAspGlyTyMetAsnGluAspAspLeuProArgSerArgArgSerArgSer 81
QY 697 GTGACCCCTTCGCGATATAATTCGCCCACTGCGAGAGAAATATACAGAGAGAGTGGAGAAC 756
Db 82 ValThrLeuProHisIleArgProValGluGluIleThrGluGluGluLeuGluAsn 101
QY 757 GTCTGCAGCAATTCGAGAGAGATATATAACCGTTCACCTGGCTCTACTGTCTCAAA 816
Db 102 ValCysSerAsnSerArgGluLysIleTyAsnArgSerLeuGlySerThrCysHisGln 121
QY 817 TGCCGTCAGAGAGACTATTGATACCAAACTGCAGAAACCCAGAGCTGTGGGGCGTT 876
Db 122 CysArgGlnLysThrIleAspThrLysThrAsnCysArgAsnProAspCysTrpGlyVal 141
QY 877 CGAGCCAGTCTGTGGCCCTGCTTCGAAACCGTTCGAGAGAGAGTTCAGGAGTGGT 936
Db 142 ArgGlyGlnPheCysGlyProCysLeuArgAsnArgTyrglyGluGluValArgAspAla 161
QY 937 CTGCTGGATCGAACTGCGCATTCGCCCTTGTGAGGAACTGCAACTGAGTTTCTGC 996
Db 162 LeuLeuAspProAsnTrpHisCysProCysArgGlyIleCysAsnCysSerPheCys 181
QY 997 CGGACGAGAGATGACCGGTGCGCACTGGGTCCTTGTGTATTATTTAGCCAAATATCATGC 1056
Db 182 ArgGlnArgAspGlyArgCysAlaThrGlyValLeuValTyLeuAlaLysTyHisGly 201
QY 1057 TTTGGGAATGTCATGCTCTACTTGAAGAGCTGAAACAGGAATTTGAATGCAAGCA 1113
Db 202 PheGlyAsnValHisAlaTyLeuLysSerLeuLysGlnGluPheGluMetGlnAla 220

RESULT 8

Q98BY9 PRELIMINARY; PRT; 207 AA.
ID Q98BY9; AC Q98BY9;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ90653.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahara K.,
RA Masuho Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK075134; BAC11425.1; -;
SQ SEQUENCE 207 AA; 23340 MW; 77AF9BB96EFD0F0F CRC64;

Alignment Scores:

Pred. No.: 3,92e-57 Length: 207
Score: 786.00 Matches: 155

Percent Similarity: 99.36% Conservative: 0
 Best Local Similarity: 99.36% Mismatches: 1
 Query Match: 40.06% Indels: 0
 DB: 2 Gaps: 0

US-10-046-935-2234 (1-1116) x Q8NBV9 (1-207)

QY 1 ATGAGCTCGCGCGTCCGAGAAAGATCTCAGAGTAAAGAACTTAAAGAAATTC 60
 DB 1 MetAspAlaArgValProGlnLysAspLeuArgValLysAsnLeuLysLysPhe 20
 QY 61 AGATATGTGAAGTTGATTTCCATCGAAACCTCGTCATCTCTGATGACAGTGTGACAGC 120
 DB 21 ArgTyrValLysLeuIleSerMetGluThrSerSerSerSerAspSerCysAspSer 40
 QY 121 TTGCTCTTCATATTTTGAACACAGAGCTCGAGTCAGTTCGGGAAGCTGTAGGACC 180
 DB 41 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr 60
 QY 181 CGCAGCAGTCGAGCAGCTCTGACCTCTCAGGCTGCGGATGAAGTTTCCAGCGCGGAGT 240
 DB 61 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 80
 QY 241 ACCAGGGGACCAACCAAAAGAGAGTCCGCGCAGCCCTCAGAGAATTCGTGACT 300
 DB 81 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 100
 QY 301 GATTCCAACTCCGATTGAGAGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 360
 DB 101 AspSerAsnSerAspSerGluAspGluSerGlyWetAsnPheLeuGluLysArgAlaLeu 120
 QY 361 AATATAAGCAAAACAAAGCAATCTCTGCAAACTCATGTCTCAATTAGAAGAGCTTCCCT 420
 DB 121 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140
 QY 421 GGCTCTGCTCGTGAAGACATCCCTCCGAGGCTCGAGCTCAGCAATCA 468
 DB 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerValSer 156

RESULT 9
 Q96C70

ID Q96C70 PRELIMINARY; PRT; 442 AA.
 AC Q96C70;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE RAM2 protein (Fragment).
 GN Name=RAM2;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]_SEQUENCE FROM N.A.
 RP SEQUENCE FROM N.A.
 RC TISSUE=Eye;
 RK MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Villalón D.K., Murzy D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalek U., Smallos D.E., Schmerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human
 RL and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Eye;
 RA Strausberg R.;
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC014630; AAH14630.2; -,
 FT NON_TER 1
 SQ SEQUENCE 442 AA; 50745 MW; C57A871694B2C4FD CRC64;

Alignment Scores:

Pred. No.: 3 4e-54 Length: 442
 Score: 752.00 Matches: 172
 Percent Similarity: 57.36% Conservative: 50
 Best Local Similarity: 44.44% Mismatches: 111
 Query Match: 38.33% Indels: 54
 DB: 2 Gaps: 10

US-10-046-935-2234 (1-1116) x Q96C70 (1-442)

QY 22 CAGAAATCTCAGAGTAAAGAACTTAAAGAAATTCAGATATGTGAAGTTGATTC 81
 DB 77 GlnSerAspLeuAsnGlyLysThrAsnProGlu-----ValMet 89
 QY 82 ATGGAACCTCGTCATCTCTGATGACAGTGTGACAGCTTCTCTGATATTTTGA 141
 DB 90 ValValGluSerAspLeuSerAspGlyLysAlaSerLeuValSerGluGluGlu 109
 QY 142 AACACGAGGCTGAGTCAGTTCGGGAAGGCTGAGGACCGCAGCCAGTCGAGCAGTCT 201
 DB 110 AspGluGlu---GluAspLysAlaThrProArgArgSerArgSerArg---ArgSerSer 127
 QY 202 GGACCTCTCAGGTGCGGATGAAGTTTCCAGCGCGGATACCAGGGGAGCAACCAACAA 261
 DB 128 IleGlyLeuArgValAlaPheGlnPheProThrLysLysLeuAlaAsnLysProAspLys 147
 QY 262 AAAGCAGAGTCCCGCCAG----- 279
 DB 148 AsnSerSerSerGluGlnLeuPheSerSerAlaArgLeuGlnAsnGluLysLysThrIle 167
 QY 280 -----CCCTCAGAGAAATCTGTGACTGATTC 306
 DB 168 LeuGluArgLysLysAspCysArgGlnValIleGlnArgGluAspSerThrSerGluSer 187
 QY 307 AACTCCGATTCAAGATCAAAAGT-----GGAATGAATTTTGGAGAAAGGGCTTTA 360
 DB 188 GluAspAspSerArgAspGluSerGlnGluSerSerAspAlaLeuLysLysArgThrMet 207
 QY 361 AATATAAGCAAAACAAAGCAATCTGCAAACTCATGTCTCAATTAGAAGCTTCCCT 420
 DB 208 AsnIleLysGluAsnLysAlaMetLeuAlaGlnLeuAlaGluLeuAsnSerMetPro 227
 QY 421 GGCTCTGCTCGTGAAGACATCCCTCCAGGCTCGAGCTCACAATCAAGGAGACCGCA 480
 DB 228 AspPhePheProValArg-----ThrProThrSerAlaSerArgLysLysThrValArg 245
 QY 481 AGCGGTACATCCCGGGTGTCTCTCCAGGAGA---AACCTGAACGGAGAGCTCTCTCT 537
 DB 246 ArgAlaPheSerGluGlyGlnIleThrArgArgMetAsnProThrArgSerAlaArgPro 265
 QY 538 CITACAGGTCAAGTCCCGGATCTCGGTCCTCTGAGCTCTACCCATGGRGAGAGCC 597
 DB 266 ProGluLysPhe-----AlaLeuGluAsnPheThrValSerAlaAla 279
 QY 598 AGTTTGGARKMTWGTACATGTTGTTGAGAAAGAGAGACCGTGGATCGCTACATGANT 657
 DB 280 LysPheAlaGluGluPheTyrSerPheArgArgArgLysThrIleGlyLysCysArg 299
 QY 658 GAAGATGACCTGCCAGAGAGCGCTCGCTCCAGATCATCGTGACCTTCCGATATAATT 717
 DB 300 Glu-----TyrArgArgHisArgIleSer-----SerPhe 310

QY	718	CGCCCAAGTGGAGAAATTACAGAGGAGGAGCTTGGAGAACCTTCGACCAATCTTCGAGAG	777
Db	311	ArgProValGluAspIleThrGluGluAspLeuGluAsnValAlaIleThrValArgAsp	330
QY	778	AAGATATATAACCGTTCACTGGGCTCTACTGTGTCATCAATGCCGTGACAGACTATTGAT	837
Db	331	LysIleTyrAspIysValLeuGlyAsnThrCysHisGlnCysArgGlnIysThrIleAsp	350
QY	838	ACCAAAACAACATGCAGAAACCCAGACTGTCTGGGCGCTTCAGAGCCCAAGTCTCTGTGGCCCC	897
Db	351	ThrIysThrValCysArgAsnGlnGlyCysCysGlyValArgGlyGlnPheCysGlyPro	370
QY	898	TGCTTTCGAAACCGTTATGTGTGAAGAGCTCAGGATGCTCTGCTGGATCGCACTGCGAT	957
Db	371	CysLeuArgAsnArgTyrGlyGluAspValArgSerAlaLeuLeuAspProAspTyrVal	390
QY	958	TGCCCGCTGTTCGAGGAATCTGCAACTGCAGATTCTTCGCCGACGACGAGATGGACGGTGT	1017
Db	391	CysProProCysArgGlyIleCysAsnCysSerTyrCysArgIysA-gaspGlyArgCys	410
QY	1018	GCGACTGGGTCCTGTGTATTATAGCAAAATATCATGGCTTTGGGAATGTCATGCGCTAC	1077
Db	411	AlaThrGlyIleLeuIleHisLeuAlaIysPheTyrGlyTyrAspAsnValLysGluTyr	430
QY	1078	TTGAAAAGCCTGAACACAGGA	1098
Db	431	LeuGluSerLeuGlnIysGlu	437
RESULT 10			
Q86YTO	ID	Q86YTO	PRELIMINARY; PRT; 453 AA.
AC	AC	Q86YTO;	
DT	01-JUN-2003	(TREMBLrel. 24, Created)	
DT	01-JUN-2003	(TREMBLrel. 24, Last sequence update)	
DT	01-OCT-2004	(TREMBLrel. 28, Last annotation update)	
DE	Transcription factor RAM2 splice variant b.		
GN	Name=RAM2;		
OS	Homo sapiens (Human).		
OC	Oco karyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
NCBI	TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Cachonen T., Wang K.H., Oswald W.B., Weitzman M.D.;		
RN	Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.		
RP	[2]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Eye;		
RC	MEDLINE=22389257; PubMed=12477932;		
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,		
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Schetz T.E.,		
RA	Brownstein M.J., Usdin T.B., Toshikiyuki S., Carninci P., Prange C.,		
RA	Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,		
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,		
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		
RA	Bradley R.W., Touchman J.W., Green E.D., Dickson M.C.,		
RA	Rodriguez A.C., Grinwald J., Schmutz J., Myers R.M., Butterfield Y.S.,		
RA	Kryzyski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,		
RA	Jones S.J., Marra M.A.;		
RT	"Generation and initial analysis of more than 15,000 full-length human		
RT	and mouse cDNA sequences."		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Eye;		

Query Match:	38.33%	Indels:	54
DB:	2	Gaps:	10
US-10-046-935-2234 (1-1116) x AAH25242 (1-453)			
Qy	22	CAGAAAGATCTCAGAGTAAGAGAACAACCTTAAGAAATTCAGATATGTGAAGTTGATTCC	81
			:::
Db	88	GlnSerAspLeuAsnGlyLysThrAsnProGlu	ValMet 100
			-----ValMet 100
Qy	82	ATGGAACCTCGTCATCTCTGATGACAGTTGTGACAGCTTTGCTTCTGATAATTTTGC	141
		:::	
Db	101	ValValGluSerAspLeuSerAspGlyLysAlaSerLeuValSerGluGluGlu	120
Qy	142	AACACAGGCTGCAGTCAGTTTCGGGAAGGCTGTAGGACCCGACGACGTGCAGGCACCTCT	201
		:::	:::
Db	121	AspGluGlu---GluAspLysAlaThrProArgArgSerArgSerArg---ArgSerSer	138
Qy	202	GGACCTCTCAGGTGGCGATGAAGTTTCCAGCGGGAGTACCAGGGGAGCAACCAACAA	261
Db	139	IleGlyLeuArgValAlaPheGlnPheProThrLysLysLeuAlaAsnLysProAspLys	158
Qy	262	AAAGCAGAGTCCCGCCAG	279

159	Db	AsnSerSerSerGluGlnLeuPheSerSerAlaArgLeuGlnAsnGluLysLysThrIle	178
280	QY	-----CCCTCAGAGAAATCTGTGACTGATCC	306
179	Db	LeuGluArgLysLysAspCysArgGlnValIleGlnArgGluaspSerThrSerGluSer	198
307	QY	AACGCCGATTCAGAAAGT-----GGAATGAATTTTGGAGAAAGGCGTTTA	360
199	Db	GluAspAspSerArgAspGluSerGlnGluSerSerAspAlaLeuLeuLysArgThrMet	218
361	QY	AATATAAGCAAAACAAAGCAATCCTTGCACAACTCTGTAATCTGAAATTAGAAAGCTTCCT	420
219	Db	AsnIleLysGluAsnLysAlaMetLeuAlaGlnLeuLeuAlaGluLeuAsnSerMetPro	238
421	QY	GGCTCGTTCCGTGGAAGACATCCCTCCAGGCTCCGACTCACAAATCAAGGAGACCGGA	480
239	Db	AspPheProValArg-----ThrProThrSerAlaSerArgLysLysThrValArg	256
481	QY	AGGGTACATTCGCGGTGTGCTTCAGAGA---AACCTGAAACGGAGAGCTCGTCT	537
257	Db	ArgAlaPheSerGluGlyGlnIleThrArgArgMetAsnProThrArgSerAlaArgPro	276
538	QY	CTTACCAGGTCACAGTCCCGATCCTCGGTCCTTGACGCTCTACCCATGGRGAGAGCC	597
277	Db	ProGluLysPhe-----AlaLeuGluAsnPheThrValSerAlaAla	290
598	QY	AGTTTGARKMTWGTACATGTCGTGAGAAAGAGAGACCGTGGATGGCTACATGAAT	657
291	Db	LysPheAlaGluGluPheTyrSerPheArgArgArgLysThrIleGlyGlyLysCysArg	310
658	QY	GAAGATGACCTGCCCAAGACCGTCGTCTCCAGATCATCCGTGACCCCTTCGGCATATAAT	717
311	Db	Glu-----TyrArgArgArgHisArgIleSer-----SerPhe	321
718	QY	CGCCACTGGAAGAAATTACAGAGGAGGATGTGGAGAACGTCGTGACGAATTTCTCGAGAG	777
322	Db	ArgProValGluAspIleThrGluAspLeuGluAsnValAlaIleThrValArgAsp	341
778	QY	AAGATATATAACCGTTCACTGGGCTCTACTGTCTCATCATCGCTCAGAAGACTATTGAT	837
342	Db	LysIleTyrAspLysValLeuGlyAsnThrCysHisGlnCysArgGlnLysThrIleAsp	361
838	QY	ACCAAAACAACTGCAGAAACCCAGACTGCTGGGCGCTTCGAGGCGCAGTCTCTGTGGCCCC	897
362	Db	ThrLysThrValCysArgAsnGlnGlyCysGlyValArgGlyGlnPheCysGlyPro	381
898	QY	TGCTTCGAACCGTTATGTGTGAAGAGGTTCAGGGATGCTCTGCTGGATCCGAAGTGGCAT	957
382	Db	CysLeuArgAsnArgTyrGlyGluAspValArgSerAlaLeuLeuAspProAspTrpVal	401

362	ILILSYSLINVALVYSLGASGSLNGLRYCYCYSGYVALATGGLYGLNINLHCYSGSLYFLD	368	
Qy	898	TGCTTCGAAACCGCTTATGGTGGAAGAGCTCAGGGATGCTCTCTGGATCCGAACATGGCAT	957
Db	382	CysLeuArgAsnArgTyrGlyGluAspValArgSerAlaLeuLeuAspProAspTrpVal	401

Query Match:	38.33%	Indels:	54
DB:	2	Gaps:	10
US-10-046-935-2234 (1-1116) x Q96GN5 (1-454)			
QY	22	CAGAAAGATCTCAGACAGTAAAGAACTTAAGAAATTCAGATATGTGAAGTTGATTTC	81
DB	89	GlnSerAspLeuAenGlyLysThrAsnProGlu-----ValMet	101
QY	82	ATGAAACCTCGTCATCTCTGATGACAGTTGTGTGACAGCTTTGCTTCTGATAAATTTGCA	141
DB	102	ValValGluSerAspLeuSerAspGlyLysAlaSerLeuValSerGluGluGluGlu	121
QY	142	AACAGAGGCTGAGTCAGTTGCGGAAGCTGTAGGACCCGACGACGTCGAGGCACTCT	201
DB	122	AspGluGlu---GluAspLysAlaThrProArgArgSerArgSerArg---ArgSerSer	139
QY	202	GGACCTCTCAGGCTGGCGATCAAGTTTCCAGCGCGAGTACCAGGGGACCAACCAAA	261
DB	140	IleGlyLeuArgValAlaPheGlnPheProThrLysLysLeuAlaAsnLysProAspLys	159
QY	262	AAACGAGATCCCGCAG-----	279
DB	160	AsnSerSerSerGluGlnLeuPheSerSerAlaArgLeuGlnAsnGluLysLysThrIle	179
QY	280	-----CCCTCAGAGAAATTTCTGTGACTGATTCC	306
DB	180	LeuGluArgLysLysAspCysArgGlnValIleGlnArgGluAspSerThrSerGluSer	199
QY	307	AATCCGATTCAGAAGATGAAGT-----GGAATGAATTTTTGGAGAAAAGGCGTTTA	360
DB	200	GluAspAspSerArgAspGluSerGlnGluSerSerAspAlaLeuLysArgThrMet	219
QY	361	AATATAAGCAAAACAAGCAATGCTTCGCAAACTCATGTCCTGAATTGAAAGCTTCCT	420
DB	220	AsnIleLysGluAsnLysAlaMetLeuAlaGlnLeuLeuAlaGluLeuAsnSerMetPro	239
QY	421	GGCTCGTTTCGTGGAAGACATCCCTCCAGGCTCCGACTCACAATCAAGGAGACCGGA	480
DB	240	AspPhePheProValArg-----ThrProThrSerAlaSerArgLysLysThrValArg	257
QY	481	AGGCGTACATTCGCGGTGTGCTTCACAGAGA---AACCTGAACGGAGAGACTCGTCT	537
DB	258	ArgAlaPheSerGluGlyGlnIleThrArgArgMetAsnProThrArgSerAlaArgPro	277
QY	538	CTTACCAGGTCACAGTCCCGATCTCGGGTCCCTTGACGCTCTACCCATCGGAGAGCG	597
DB	278	ProGluLysPhe-----AlaLeuGluAsnPheThrValSerAlaAla	291
QY	598	AGTTTGARKMTMGTCATGTTGGTGAGAAGAGAGAACCGTGATGGCTACATGAAT	657
DB	292	LysPheAlaGluGluPheTyrSerPheArgArgArgLysThrIleGlyLysCysArg	311
QY	658	GAAGATGACCTGCCAGAAAGCGTCGCTCCAGATCATCCGTGACCTTCGCGATATAAT	717
DB	312	Glu-----TyrArgArgGluHisArgIleSer-----SerPhe	322
QY	718	CGCCAGTGGAGAATAATACAGAGAGAGGTGGAGAACGTCCTGCAGCAATTTCTCGAGAG	777
DB	323	ArgProValGluAspIleThrGluAspLeuGluAsnValAlaIleThrValArgAsp	342
QY	778	AAGATATATACCGTTCACTGGGCTCTACTTGTCTCATCAATGCCGTCAGAGACTATTGAT	837
DB	343	LysIleTyrAspLysValLeuGlyAsnThrCysHisGlnCysArgGlnLysThrIleAsp	362
QY	838	ACCAAAACAACTCGACAAACCCAGACTGCTGGGCGCTTCAGAGGCGAGTTCTGTGCGCCC	897
DB	363	ThrLysThrValCysArgAsnGlnGlyCysCysGlyValArgGlyGlnPheCysGlyPro	382
QY	898	TGCCTTCGAACCGTTATGTTGAAGAGTCAAGGATGCTCTGCTGGATCCGAACTGGCAT	957
DB	383	CysLeuArgAsnArgTyrGlyGluAspValArgSerAlaLeuLeuAspProAspTrpVal	402

101	ValValGluSerAspLeuSerAspAspGlyIysAlaSerLeuValSerGluGluGlu	120
142	AAACACGAGGCTGCAGTCCAGTTCGGGAGGCTGTAGGACCCGACGAGTCGACGACCTCT	201
121	AspGluGlu---GluAspIysAlaThrProArgSerArgSerArg---ArgSerSer	138
202	GGACCTCTCAGGTGGCGATGAAGTTTCCAGCGCGGAGTACCAGGGGAGCAACCAAAA	261
139	IleGlyLeuArgValAlaPheGlnPheProThrLysIysLeuAlaAsnLysProAspLys	158
262	AAAGCAGAGTCCCGCCAG-----	279
159	AsnSerSerGluGlnLeuPheSerSerAlaArgLeuGlnAsnGluLysLysThrIle	178
280	-----CCCTCAGAGAAATCTGTGACTGATTC	306
179	LeuGluArgLysIysAspCysArgGlnValIleGlnArgGluAspSerThrSerGluSer	198
307	AACTCCGATTACGAAGATGAAGT-----GGAATGAATTTTTTGGAGAAAGGGCTTTA	360
199	GluAspAspSerArgAspGluSerGlnGluSerSerAspAlaLeuLeuLysArgThrMet	218
361	AATATAAGCAAAACAAAGCAATCTTGTGAAAACACTCATGTCTGAATAGAAAGCTTCCCT	420
219	AsnIleLysGlnAsnLysAlaMetLeuAlaGlnLeuAlaGluLeuAsnSerMetPro	238
421	GGCTCGTTCCGTGGAGACATCCCTCCCAAGGCTCCGACTCACAAATCAAGGACACCGCA	480
239	AspPhePheProValArg-----ThrProThrSerAlaSerArgLysLysThrValArg	256
481	AGGCGTCATCTCCGGGTGTGTCTCCAGGAGA---AACCCCTGAACGGAGAGCTCGTCCT	537
257	ArgAlaPheSerGluGlyGlnIleThrArgArgMetAsnProThrArgSerAlaArgPro	276
538	CTTACACAGTCAAGTCCCGGATCTCGGGTCCCTTGACGCTCTACCCATGGRGAGAGCC	597
277	ProGluLysPhe-----AlaLeuGluAsnPheThrValSerAlaA	290
598	AGTTTGARKMTWGTACATGTTGGTCAGAAAGAGGAAGACCTGGATGGCTACATCAAT	657
291	LysPheAlaGluPheThrSerPheArgArgArgLysThrIleGlyLysCysArg	310
658	GAAGATGACCTGCCAGAGACCTCGCTCCAGATCATCCGTGACCCCTTCGCGCATATAAT	717
311	Glu-----TyrArgArgArgHisArgIleSer-----SerPhe	321
718	CGCCGAGTGAAGAAATTACAGAGGAGGTGGAGAACGCTGTCAGACAATTCGAGAG	777
322	ArgProValGluAspIleThrGluGluAspLeuGluAsnValAlaIleThrValArgAsp	341
778	AAGATATATAACCGTTCACTGGGCTCTACTGTGCATCAATCGCTCAGAGACTATTGAT	837
342	LysIleCysAspLysValLeuGlyAsnThrCysHisGlnCysArgGlnLysThrIleAsp	361
838	ACCAAAACAAATTCGAGAAACCCAGACTGCTGGGGGGCTTCGAGGCCAGTTCTGTGCCCC	897
362	ThrLysThrValCysArgAsnGlnGlyCysCysGlyValArgGlyGlnPheCysGlyPro	381
898	TGCCTTCGAAACCGTTATGGTGAAGAGGTGAGGATGCTCTGCTGATCCGAACCTGGCAT	957
382	CysLeuArgAsnArgTyrGlyGluAspValArgSerAlaLeuLeuAspProAspTyrVal	401
958	TGCCCGCTTGTGAGGAATCTGCAACTCGATTCTGCGCGCAGCAGATCGACCGTGT	1017
402	CysProProCysArgGlyIleCysAsnCysSerTyrCysArgLysArgAspGlyArgCys	421
1018	GCAGCTGGGGTCTTGTGATTTTAGCCAAATATCATGGCTTTGGGAATGTGCATGCTAC	1077
422	AlaThrGlyIleLeuIleHisIleuAlaLysPheTyrGlyTyrAspAsnValLysGluTyr	441
1078	TTGAAAGCGCTGAACACAGAA	1098
442	LeuGluSerLeuGlnLysGlu	448


```

RESULT 14
RAH32576
ID AAH32576 PRELIMINARY; PRT; 453 AA.
AC AAH32576;
DT 29-MAR-2004 (TrEMBLrel. 27, Created)
DT 29-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 29-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Transcription factor RAM2.
GN RAM2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Rubin A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullane P.H.,
RA Boeak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.W., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC032576; AAH32576.1; -.
SQ SEQUENCE 453 AA; 52018 MW; 6AA046560F31DABB CRC64;

Alignment Scores:
Pred. No.: 1,95e-53 Length: 453
Score: 743.00 Matches: 171
Percent Similarity: 57.11% Conservative: 50
Best Local Similarity: 44.19% Mismatches: 112
Query Match: 37.87% Indels: 54
DB: 2 Gaps: 10

US-10-046-935-2234 (1-1116) x AAH32576 (1-453)
QY 22 CAGAAAGATCTCAGAGTAAGAAAGAACTTAAAGAAATTCAGATATGTGAAGTTGATTTC 81
Db 88 GlnSerAspLeuAsnGlyLysThrAsnProGlu-----ValMet 100
QY 82 ATGGAACCTCTGTCATCTCTGATGACAGTTGTGACAGTTGTGCTGATATTTTGA 141
Db 101 ValValGluSerAspLeuSerAspGlyLysAlaSerLeuValSerGluGluGlu 120
QY 142 AACACGAGGCTCAGTTCAGTTCCGGAAGCTGTAGGACCGGACGAGTGCAGGACTCT 201
Db 121 AspGluGlu---GluAspLysAlaThrProArgArgSerArg---ArgSerSer 138
QY 202 GCACCTCTTCAGGTCGCGATGAAGTTTCCAGGCGGAGTACACGAGGAGCAACCAAA 261
Db 139 IleGlyLeuArgValAlaPheGlnPheProThrLysLeuAlaAsnLysProAspLys 158
QY 262 AAAGCAGAGTCCCGCCAG----- 279

```


Blank

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: October 12, 2004, 15:04:25 ; Search time 112 Seconds
(without alignments)
7148.962 Million cell updates/sec

Title: US-10-046-935-2234
Perfect score: 1962
Sequence: 1 atggacgtcgcgcggtgcc.....aatttgaaatgaacataa 1116

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 4004546

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-O=/cpn2_1/USPTO spool_p/US10046935/runat_12102004_155225_15233/app_query.fasta_1.1287
-DB=A_Geneseq_23Sep04 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITIS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10046935 @CGN 1.1 1.77 @runat_12102004_155225_15233 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_23Sep04:.*
1: Geneseq1980s:.*
2: Geneseq1990s:.*
3: Geneseq2000s:.*
4: Geneseq2001s:.*
5: Geneseq2002s:.*
6: Geneseq2003as:.*
7: Geneseq2003bs:.*
8: Geneseq2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1943	99.0	371	5	Abb05376 Human col
2	1929	98.3	371	4	Aab94473 Human pro
3	1928	98.3	371	6	Abu56624 Lung canc
4	1929	98.3	371	7	Adf76612 Novel hum
5	1929	98.3	371	7	Adn38808 Human CDC
6	1929	98.3	371	8	Adg09178 Human CDC
7	1929	98.3	407	3	Aab43720 Human can
8	1929	98.3	407	4	Aag74880 Human col
9	1924	98.1	397	6	Ada98284 Human sec
10	1924	98.1	397	6	Ada44119 Human sec

11	1924	98.1	397	7	ADC20449	Adc20449 Human sec
12	1924	98.1	397	7	ADF10748	Adf10748 Human sec
13	1879.5	95.8	450	4	AAB93122	Aab93122 Human pro
14	1879.5	95.8	450	8	ADL91540	Adl91540 Human imm
15	1879.5	95.8	450	8	ADN06070	Adn06070 Antipsori
16	1879.5	95.8	450	8	ADO20297	Ado20297 Human PRO
17	1879.5	95.8	450	8	ADQ09208	Adq09208 Human CDC
18	1879.5	95.8	450	8	ADQ09176	Adq09176 Human CDC
19	1656	84.4	347	5	AAG79357	Aag79357 CGA7 pref
20	1656	84.4	347	7	ADN39876	Adn39876 Cancer/an
21	1555	79.3	298	8	ADN99718	Adn99718 Novel hum
22	1520.5	77.5	320	3	AAB23164	Aab23164 Human col
23	1520.5	77.5	320	5	AAG79356	Aag79356 CGA7. 8/2
24	1520.5	77.5	320	7	AD848023	Ad848023 Human CGA
25	852	43.4	174	6	ABU71082	Abu71082 Human adi
26	786	40.1	207	4	AAM93560	Aam93560 Human pol
27	786	40.1	207	8	ADL31298	Adl31298 Human pro
28	752	38.3	420	7	ADM04753	Adm04753 Human pro
29	563.5	28.7	110	4	ABG22773	Abg22773 Novel hum
30	535	27.3	89	3	AAB32062	Aab32062 Human sec
31	498.5	25.4	222	4	ABG03573	Abg03573 Novel hum
32	424.5	21.4	344	4	ABG03571	Abg03571 Novel hum
33	346	17.6	71	4	ABG22772	Abg22772 Novel hum
34	241.5	12.3	102	3	AAB32061	Aab32061 Human sec
35	195	9.9	40	3	AAB32010	Aab32010 Human sec
36	171.5	8.7	268	4	AAB94363	Aab94363 Human pro
37	146	7.4	875	8	ADD062965	Add062965 Transcript
38	138	7.0	786	7	ADD30845	Add30845 Plant yie
39	138	7.0	786	8	ADI43825	Adi43825 Plant tra
40	132.5	6.8	570	8	ADN72417	Adn72417 Thale cre
41	131.5	6.7	906	8	ADO61723	Ado61723 Transcript
42	129	6.6	754	5	AAB24596	Aae24596 Human SR-
43	129	6.6	754	7	ADG10664	Adg10664 Human STA
44	129	6.6	754	8	ADH09504	Adh09504 Human hos
45	127	6.5	448	7	ADG10780	Adg10780 Human STA

ALIGNMENTS

RESULT 1
ABB05376
ID ABB05376 standard; protein; 371 AA.
XX
AC ABB05376;
XX
DT 08-APR-2002 (first entry)
XX
DE Human colon tumour antigen polypeptide SEQ ID NO:2235.
XX
KW Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
XX
XX colon tumour metastatic antigen; diagnosis.
OS Homo sapiens.
XX
XX WO200196388-A2.
XX
XX 20-DEC-2001.
XX
XX 08-JUN-2001; 2001WO-US018557.
XX
XX 09-JUN-2000; 2000US-0210899P.
XX
XX 20-FEB-2001; 2001US-0270216P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Jiang Y, Harlocker SL, Secrist H;
XX
XX WPI; 2002-114514/15.
XX
XX Novel isolated colon tumor polynucleotide differentially expressed in
XX
XX colon tumor or colon metastatic tumor and polypeptides encoded by them,
XX
XX useful for inhibiting development of cancer in patient.

PS Claim 2; SEQ ID NO 2235; 105pp; English.

XX ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (I)

CC which were isolated from human colon tumour and colon metastatic tumour

CC cDNA libraries. (I) have cyrostatic activity and can be used in vaccine

CC production. (I) can be used for stimulating and/or expanding T cells

CC specific for a tumour protein on contact with the T cells. They are also

CC useful for inhibiting the development of cancer in a patient. (I) can be

CC used as probes or primers for nucleic acid hybridisation, for preparing

CC mutant species primers, or primers for use in genetic constructions. (I)

CC can be used in the diagnosis of a colon tumour. The present sequence

CC represents a human colon tumour antigen amino acid sequence which is

CC specifically claimed in the present invention

XX Sequence 371 AA;

Alignment Scores:

Pred. No.:	Score:	6,078-196	Length:	371
Percent Similarity:	1943.00	Matches:	371	
Best Local Similarity:	100.00%	Conservative:	0	
Query Match:	99.03%	Mismatches:	0	
DB:	5	Indels:	0	
		Gaps:	0	

US-10-046-935-2234 (1-1116) x ABB05376 (1-371)

QY 1 ATGGAGCTCGCGGTGCGGAGAAAGTCTCAGAGTAAGAAAGAACTTAAAGAAATTC 60

DB 1 MetAspAlaArgValProGlnLysAspLeuArgValLysAsnLeuLysPhe 20

QY 61 AGATATGTGAAGTTGATTTCCATGGAAACCTCGTCATCTCTGATGACAGTTGTGACAC 120

DB 21 ArgTyrValLysLeuIleSerMetGluThrSerSerSerSerSerSerSerSerSerSer 40

QY 121 TTTGCTTCTGATAATTTTCCAAACACAGAGCTGCAGTTCAGTTCGGGAAGCTGTAGGAC 180

DB 41 PheAlaSerAspAsnPheAlaSerThrArgLeuGlnSerValArgGluGlyCysArgThr 60

QY 181 CGCAGCAGTGCAGGCACTCTGGACCTCTCAGGTGGCGATGAAAGTTTCCAGCGCGAGCT 240

DB 61 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 80

QY 241 ACCAGGGGACCAACCAACAAAGCAGAGTCCCGCCAGCCCTCAGAGAATTTCTGTGACT 300

DB 81 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 100

QY 301 GATTCACATCCGATTCAGAGATGAAGTGAATGAATTTTTCGAGAAAGGGCTTTA 360

DB 101 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 120

QY 361 AATATAAGCAAAACAAAGCAATGCTTGCAGAACTCATGCTGAATTAGAAGCTTCCCT 420

DB 121 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140

QY 421 GGCCTGCTTCGTGGAGACATCCCTCCAGGCTCCGACTCACAATCAAGAGGACCGCGA 480

DB 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 160

QY 481 AGCGGTACATTCGCGGTGCTGCTCCAGGAGAAACCTTGACGGAGAGCTGCTCCTTT 540

DB 161 ArgArgThrPheProGlyValAlaSerArgAsnProGluArgArgAlaArgProLeu 180

QY 541 ACCAGGTCAAGTTCGCGGTCTCGGGTCCCTTGACGCTCTACCCATGGGAGAGCAGT 600

DB 181 ThrArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMet***ArgAlaSer 200

QY 601 TTKGARKMTWGPACATGTTGGTGAGAAAGAGAGAACCGTGGATGGCTACATGAATGA 660

DB 201 ***Glu*****TyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 220

QY 661 GATGACCTGCCAGAGACCGTCCCTCCAGATCATCGTGCACCTTCCGGCATATAATTCG 720

DB 221 AspAspLeuProArgSerArgArgSerArgSerValThrLeuProHisIleLeuArg 240

QY 721 CCAGTGGAGAAATTTACAGAGGAGGAGTTCGAGAAAGCTCTGAGCAATTTCTGAGAGAAG 780

DB 241 ProValGluGluIleThrGluGluLeuGluAsnValCysSerAsnSerArgGluLys 260

QY 781 APATATAACCGTTCCTCTGGCTCTACTTGTCTCATCAATGCGCTCAGAGAAGCTATTGATACC 840

DB 261 IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 280

QY 841 AAAACAAACTGCAGAAACCCAGACTGCTGGGCGCTTCGAGGCCAGTCTCTGTCGCCCTCG 900

DB 281 LysThrAsnCysArgAsnProAspCysTrpGlyValArgGlyGlnPheCysGlyProCys 300

QY 901 CTTGAAACCGTTATGCTGAGAGGTCAGAGGATGCTCTGCTGGATCCGAACTGGCATTCG 960

DB 301 LeuArgAsnArgTyrGlyGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 320

QY 961 CCGCTTGTGAGGAATCTGAACTGCACTTCTGCGCGCAGCAGATGAGCGTGTGGG 1020

DB 321 ProProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 340

QY 1021 ACTGGGCTCTGCTGTATTAGCAAAATATCATGCTTTGGGAATGCTGATGCTACTTGG 1080

DB 341 ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 360

QY 1081 AAAAGCCTGAAACAGCAATTTGAAATGCAAGCA 1113

DB 361 LysSerLeuLysGlnGluPheGluMetGlnAla 371

RESULT 2

AAB94473

ID AAB94473 standard; protein; 371 AA.

XX AAB94473;

XX 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:15139.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-00116126.

XX 29-JUL-1999; 99JP-00248036.

XX 27-AUG-1999; 99JP-00300253.

XX 11-JAN-2000; 2000JP-00118776.

XX 02-MAY-2000; 2000JP-00183767.

XX 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

XX Claim 8; SEQ ID NO 15139; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602

PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer.

XX PS Claim 27; Page 353; 453pp; English.

XX CC The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridizes
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung
 CC cancer in a patient and for treating a mammal having lung cancer by
 CC administering a modulatory compound identified. The methods are useful
 CC for treating lung cancer, such as small cell lung cancer, non-small cell
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
 CC for diagnostic purposes and as targets for screening for therapeutic
 CC compounds that modulate lung cancer, such as antibodies. Sequences
 CC ABUS6408-ABUS6745 represent lung cancer-associated polypeptides of the
 CC invention

XX SQ Sequence 371 AA;

Alignment Scores:

Pred. No.: 1,84e-194 Length: 371
 Score: 1929.00 Matches: 364
 Percent Similarity: 98.11% Conservative: 0
 Best Local Similarity: 98.11% Mismatches: 7
 Query Match: 98.32% Indels: 0
 DB: 6 Gaps: 0

US-10-046-935-2234 (1-1116) x ABUS6624 (1-371)

QY 1 ATGGAGCGCTCGCGCGCGGAGAAAGATCTCAGAGTAAAGAAAGAAATTC 60
 DB 1 MetAspAlaArgValProGlnLysAspLeuArgValLysLysAsnLeuLysPhe 20
 QY 61 AGATATGTGAAGTTGATTTCCATGGAACCTGTCATCTCGATGACAGTTGTGACAGC 120
 DB 21 ArgTyrValLysLeuLysSerMetGluThrSerSerSerSerSerSerSerSerSer 40
 QY 121 TTTGCTTCTGATATTTTCAACACAGAGCTGCAGTCAAGTTCGGGAAGGCTGTAGGAC 180
 DB 41 PheAlaSerAspAsnPheAlaSerThrArgLeuGlnSerValArgGluGlyCysArgThr 60
 QY 181 CGCAGCGCAGTGCAGGCACTCTGGACCTCTCAGGTCGCGATGAAGTTTCCAGCGCGAGT 240
 DB 61 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 80
 QY 241 ACCAGGGGAGCAACCAACAAAAGCAGAGTCCCGCCAGCCCTCAGAGAATTTCTGTGACT 300
 DB 81 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 100
 QY 301 GATTTCCAATCCGATTCAGAGATGAAGTGAATGAATTTTTCAGAGAAGGGCTTTA 360
 DB 101 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGlnLysArgAlaLeu 120
 QY 361 AATATAAGCAAAACAAAGCAATGCTTGCAAAACCTCATGCTGAATTAGAAGCTTCCCT 420
 DB 121 AsnLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140
 QY 421 GGCTGCTTCGTTGGAGACATCCCTCCAGGCTCCGACTCACAATCAAGGACCGCA 480
 DB 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 160
 QY 481 AGCGGTACATTCCTCCGGTGTGTCTCCAGGAGAAACCTGAAACGAGAGCTCTCTCTT 540
 DB 161 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgAlaArgProLeu 180

QY 541 ACCAGGTCAAGTCCCGATCCTCGGTCCCTTACCGCTCTACCCATGGAGAGCCAGT 600
 DB 181 ThrArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMetGluGluGlu 200
 QY 601 TTKGARKMTMTGATCATGTTGTGTGAGAAAGAGAGACCGTGGATGGCTACATGAATGAA 660
 DB 201 GluGluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 220
 QY 661 GATGACCTGCCAGAGCGCTCGCTCCAGATCATCGTGACCTCCGCTCCGATATTAATTCG 720
 DB 221 AspAspLeuProArgSerArgSerArgSerValThrLeuProHisIleIleArg 240
 QY 721 CCAGTGGAGAAATTACAGAGGAGGAGTGGAGAAAGCTCTGCAGCAATTTCTCGAGAGAAG 780
 DB 241 ProValGluGluIleThrGluGluLeuGluAsnValCysSerAsnSerArgGluLys 260
 QY 781 ATATATAACCGTTCATCTGGCTCTACTTGTCTCATCAATGCGCTCAGAAGACTATTGATACC 840
 DB 261 IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 280
 QY 841 AAAACAAACTGCAGAAACCCAGACTGCTGGGCGGTTCCAGAGGCGAGTTCTGTGCCCCCTGC 900
 DB 281 LysThrAsnCysArgAsnProAspCysTrpGlyValArgGlyGlnPheCysGlyProCys 300
 QY 901 CTTGAAACCGTTATGTGTGAAGAGGTTCAGGGATGCTCTCTGATCCGAACTGGCATTCG 960
 DB 301 LeuArgAsnArgTyrGlyGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 320
 QY 961 CCGCTTGTGAGGAATCTGAACTGCAGTCTTCTGCCGCGCAGGAGATGCGGTGTGG 1020
 DB 321 ProProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 340
 QY 1021 ACTGGGCTCTGTGTATTAGCCAAATATCATGGCTTTGGGAATGTGATGCTACTTGTG 1080
 DB 341 ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 360

RESULT 4

ADF76612

ID ADF76612 standard; protein; 371 AA.

XX ADF76612;

AC ADF76612;

XX 26-FEB-2004 (first entry)

DT 26-FEB-2004 (first entry)

XX Novel human secreted and transmembrane protein SeqID.286.

XX human; PRO; membrane bound protein; membrane bound receptor;
 KW cell proliferation; cell migration; cell differentiation;
 KW mitogenic factor; survival factor; cytotoxic factor;
 KW differentiation factor; neuropeptide; hormone; cell receptor;
 KW receptor-ligand interaction; cytostatic; chondrocyte; tumour.

XX Homo sapiens.

XX WO2003072035-A2.

XX 04-SEP-2003.

XX 21-FEB-2003; 2003WO-US005241.

XX 22-FEB-2002; 2002US-0359461P.

XX (GETH) GENENTECH INC.

XX Bodary SC, Clark H, Hunte B, Jackman JK, Schoenfeld JR;
 PI Williams PM, Wood WI, Wu TD;
 XX WPI; 2003-721702/68.

DR N-PSDB; ADF76611.
XX
PT New PRO polypeptides, useful for diagnosing and treating an immune
PT related disorder, e.g. systemic lupus erythematosus, rheumatoid
PT arthritis, osteoarthritis, juvenile chronic arthritis, thyroiditis or
PT diabetes mellitus.
XX
PS Claim 10; SEQ ID NO 286; 918pp; English.
XX
CC This invention relates to novel nucleic acids encoding human PRO secreted
CC and transmembrane proteins. Extracellular proteins play important roles
CC in the formation, differentiation and maintenance of multicellular
CC organisms. The fate of many individual cells (for example proliferation,
CC migration or differentiation) is typically governed by information
CC received from other cells and the immediate environment. The information
CC is often transmitted by secreted polypeptides (for example mitogenic
CC factors, survival factors, cytotoxic factors, differentiation factors,
CC neurotrophins and hormones) which are received and interpreted by diverse
CC cell receptors or membrane bound proteins. These membrane bound proteins
CC and receptors may be of use as pharmaceutical and diagnostic agents, such
CC as in the blocking of receptor-ligand interactions. The current invention
CC provides the amino acid sequences of novel human membrane bound receptors
CC and proteins, along with the cDNA sequences encoding them. The novel
CC proteins of the invention may have cytostatic activities through the
CC stimulation of chondrocytes. The nucleic acids of the invention may be
CC useful for the manufacture of a medicament for diagnosing or treating a
CC tumour in a mammal. In addition, they may be useful for measuring or
CC detecting the expression of a tumour associated gene. The present
CC sequence is the amino acid sequence of a human PRO protein of the
CC invention.
XX
SQ Sequence 371 AA;
Alignment Scores:
Pred. No.: 1,84e-194 Length: 371
Score: 1929.00 Matches: 364
Percent Similarity: 98.11% Conservative: 0
Best Local Similarity: 98.11% Mismatches: 7
Query Match: 98.32% Indels: 0
DB: 7 Gaps: 0
US-10-046-935-2234 (1-1116) x ADF76612 (1-371)
QY 1 ATCGAGCTCCGCGTCCGCGAGAAAGATCTCAGAGTAAAGAGAACTTAAAGAAATTC 60
Db 1 MetAspAlaArgValProGlnLysAspLeuArgValLysAsnLeuLysLysPhe 20
QY 61 AGATATGTGAAGTTGATTTCCATGGAACCTCGTCATCCTCTGATGACAGTGTGACGC 120
Db 21 ArgTyrValLysLeuLysSerMetGluThrSerSerSerSerSerSerSerSerSer 40
QY 121 TTTGCTTCTGATAATTTTGCAACACGAGGCTGCAGTCAGTCAGTCGGGAGGCTGAGACC 180
Db 41 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr 60
QY 181 CGCAGCAGTCGAGGCACTCTGGACCTCTCAGGCTGGCGATGAAGTTTCAGCGCGGAGT 240
Db 61 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 80
QY 241 ACCAGGGGACCAACCAAAAGCAGAGTCCGCCAGCCCTCAGAGAATTCGTGACT 300
Db 81 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 100
QY 301 GATTCCAACCTCCGATTCAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 360
Db 101 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysAlaLeu 120
QY 361 AATATAAGCAAAACAAAGCAATGCTTGCAGAACTCATGTCTGAATAGAAAGTCCCT 420
Db 121 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140
QY 421 GCGTCGTTCCGTGGAGACATCCCTCCAGGCTCCGACTCAAAATCAAGGACCGCGA 480

Db 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 160
QY 481 AGCGGTACATTCGCGGTGTTGCTTCAGGAGAAACCTGAACGAGAGCTGCTCTT 540
Db 161 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgArgAlaArgProLeu 180
QY 541 ACCAGGTCAAAGTCCCGGATCCCTCGGCTCCCTTCAACCTCTTACCCTACCGAGAGCCAGT 600
Db 181 ThrArgSerArgSerArgLysLeuGlySerLeuAspAlaLeuProMetGluGluGlu 200
QY 601 TTKGARKMTVMGTACATGTTGGTGAGAAAGAGGAGGAGCCGTGGATGGCTACATGAATCAA 660
Db 201 GluGluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 220
QY 661 GATGACCTGCCAGAGCCGCTCGCTCCAGATCATCCGTGACCTCCGTCATATAATTCGC 720
Db 221 AspAspLeuProArgSerArgSerArgSerValThrLeuProHisIleArg 240
QY 721 CCAGTGGAAGAAATTCAGAGGAGGAGTTCGAGAACCTCTGCAGCAATTCCTCGAGAGAAG 780
Db 241 ProValGluGluLeuThrGluGluGluLeuAsnValCysSerAsnSerArgGluLys 260
QY 781 ATATATAACCGTTCACGTGGCTCTACTGTGTCATCAATGCCGTCAAGAAGACTATTGATACC 840
Db 261 IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 280
QY 841 AAAACAACCTGCAGAACCCAGACTGCTGGGGCGTTCGAGCCAGTTCCTGTGGCCCTGC 900
Db 281 LysThrAsnCysArgAsnProAspCysTyrPglYValArgGlyGlnPheCysGlyProCys 300
QY 901 CTTTCGAACCGTTCATGTTGAGAGGTCAGGAGTCTGCTGGATCCGAATCGCATTCGCATTGC 960
Db 301 LeuArgAsnArgTyrGlyGluGluValArgAspAlaLeuLeuAspProAsnThrHisCys 320
QY 961 CCGCTTGTTCGAGAAATTCGAACTGCACTGAGTTTCGCGGAGCAGATGACCGTGTGCG 1020
Db 321 ProProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 340
QY 1021 ACTGGGGTCTCTGTTATTTAGCCAAATATCATGCTTCGGTTCGGAAATGTGCATGCTACTTG 1080
Db 341 ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 360
QY 1081 AAAAGCTCAACAGAGAAATTTGAAATGCAAGCA 1113
Db 361 LysSerLeuLysGlnGluPheGluMetGlnAla 371
RESULT 5
ADN38808
ID ADN38808 standard; protein; 371 AA.
XX
AC ADN38808;
XX
XX 17-JUN-2004 (first entry)
DT
DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:126.
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; lechaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnary; gene therapy; vaccine.
XX
OS Homo sapiens.
XX
XX WO2003042661-A2.
XX
XX 22-MAY-2003.
XX
XX 13-NOV-2002; 2002WO-US036810.
PF
XX 13-NOV-2001; 2001US-0350666P.
XX

Db 261 IleTyrAsnArgSerIeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 280
 QY 841 AAAACAACTGCAGAAACCCAGACTGCTGGGGGTTCGAGGCCAGTTCGTGGCCCTGC 900
 Db 281 LysThrAsnCysArgAsnProAspCysTrpGlyValArgGlyGlnPheCysGlyProCys 300
 QY 901 CTTCCGAACCGTTATGTTGAAGAGTTCAGGGATGCTCTCTGTGATCCGAACCTGGCATTC 960
 Db 301 LysArgAsnArgTyrGlyGluGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 320
 QY 961 CGCCTTGTCCAGGAATCTGCAACTGCTTCCTGCGCAGGAGATGCGAGTGTGG 1020
 Db 321 ProProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 340
 QY 1021 ACTGGGCTCTCTGTGATTTAGCAATATCATGGCTTTGGGAATGTGATGCTACTTGG 1080
 Db 341 ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 360
 QY 1081 AAAAGCCTGAACAGCAATTTGAAATGCAAGCA 1113
 Db 361 LysSerLeuLysGlnGluPheGluMetGlnAla 371

RESULT 7

AAB43720

ID AAB43720 standard; protein; 407 AA.

XX AAB43720;

XX 08-FEB-2001 (first entry)

XX Human cancer associated protein sequence SEQ ID NO:1165.

XX Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
 KW antidiabetic; antiaslthmatic; antirheumatic; antiarthritic; antiviral;
 KW antinflammatory; antithyroid; antiallergic; antibacterial; cardiac;
 KW dermatological; neuroprotective; thrombolytic; coagulant; neutropic;
 KW vasotropic; antiporatic; antidiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening.

XX Homo sapiens.

XX WO200055350-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US005882.

XX 12-MAR-1999; 99US-0124270P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2000-587533/55.

XX N-PSDB; AAC77929.

PT Novel isolated nucleic acids comprising sequences encoding peptides
 PT useful for treating or diagnosing e.g. cancer.

XX Claim 11; Page 1785-1787; 2352pp; English.

XX AAC77607 to AAC78448 encode the human cancer associated proteins given in
 CC ABA43398 to ABA44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnery; immunomodulator;
 CC antidiabetic; antiaslthmatic; antirheumatic; antiarthritic;
 CC antinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
 CC dermatological; neuroprotective; cardiac; thrombolytic; coagulant;
 CC neutropic; vasotropic; antiporatic and antiangiogenic. The

CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and ABA44240 represent sequences used in the exemplification of
 CC the present invention

XX Sequence 407 AA;

Alignment Scores:
 Pred. No.: 1,92e-194 Length: 407
 Score: 1929.00 Matches: 364
 Percent Similarity: 98.11% Conservative: 0
 Best Local Similarity: 98.11% Mismatches: 7
 Query Match: 98.32% Indels: 0
 DB: Gaps: 0

US-10-046-935-2234 (1-1116) x AAB43720 (1-407)

QY 1 ATGACGCTCGCGCTGCCGAGAAAGATCTCAGATGAAGAAAGAACTTAAAGAATTC 60

Db 37 MetAspAlaArgArgValProGlnLysAspLeuArgValLysAsnLeuLysPhe 56

QY 61 AGATATGTGAAGTTGATTTCCATGGAACCTCGTCATCTCTCTGATCAGATGTCGACGC 120

Db 57 ArgTyrValLysLeuIleSerMetGluThrSerSerSerSerSerSerSerSerSer 76

QY 121 TTGCTCTCTGATAATTTTGCAACACGAGCTCGCTCAGTTCGGAAGGCTGTAGAAC 180

Db 77 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr 96

QY 181 CGCAGCAGTGCAGGCACTCTGCACTCTCAGGCTGGCGATGAAGTTTCAGCGCGAGT 240

Db 97 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 116

QY 241 ACCAGGGAGCAACCAACAAAGAGAGTCCGCCGCCCTCAGAGATTCGTGACT 300

Db 117 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 136

QY 301 GATTCCAACTCCGATTCAGAAAGATGAAGTGAATGAATTTTGGAGAAAAGGCTTTA 360

Db 137 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 156

QY 361 AATATAAGCAAAACAAAGCAATGCTTGAACAACTCATGTCTGAAATAGAAAGTTCCT 420

Db 157 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 176

QY 421 GGCTCGTTCGTCGAGAGACATCCCTCCAGGCTCCGACTCACAATCAAGAGACCGCGA 480

Db 177 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 196

QY 481 AGGCGTACATTCCTGGGTGTGCTTCAGGAGAAAACCTCAAGCGAGAGCTCGTCTCTT 540

Db 197 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgArgAlaArgProLeu 216

QY 541 ACCAGGTCAAGTCCCGGATTCCTCGGTCCTTTCAGCTCTACCCATGCGAGAGCCAGT 600

Db 217 ThrArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMetGluGluGlu 236

QY 601 TTGARKMTWGTACATGTTGGTGAAGAAAGAGAAAGCCGTGGATGGCTACATGAATGAA 660

Db 237 GluGluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 256

QY 661 GATGACCTGCCAGAGCCCTCGCTCCAGATCATCCGTGACCCCTCCGCAATATAATTCGC 720

Db 257 AspAspLeuProArgSerArgArgSerValThrLeuProHisIleArg 276

QY 721 CCAGTGAAGAAATACAGAGGAGCTTGAGAACGCTCTGCAGCAATTTCTCGAGAGAAG 780
DB 277 ProValGluGluThrGluGluGluValCysSerAsnSerArgGluLys 296
QY 781 ATATATACCGTTCACTGGGCTCTACTTGTGCATCAATGCCGTCAGAGACTATTGATACC 840
DB 297 IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 316
QY 841 AAAACAACATGCAGAACCCAGACTGCTGGGCGCTTCGAGCCAGTTCTGTGGCCCTGC 900
DB 317 LysThrAsnCysArgAsnProAspCysTrpGlyValArgGlyGlnPheCysGlyProCys 336
QY 901 CTTGGAACCGTTATGTGTGAAGAGCTCAGGATCTCTGCTGGATCCGAATGCACTTGC 960
DB 337 LeuArgAsnArgTyrGlyGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 356
QY 961 CCGCTTGTGAGGAATCTGCAACTGCACTGTTCTGCGGAGGAGATGACGCTGTGCG 1020
DB 357 ProProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 376
QY 1021 ACTGGGCTCTCTGTATTTAGCCAAATATCATGCTTGGGATGTGCATGCTACTTG 1080
DB 377 ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 396
QY 1081 AAAAGCCTGAACAGGAATTTGAAATGCAAGCA 1113
DB 397 LysSerLeuLysGlnGluPheGluMetGlnAla 407

RESULT 8

AAG74880

ID AAG74880 standard; protein; 407 AA.

AC AAG74880;

XX 03-SEP-2001 (first entry)

XX Human colon cancer antigen protein SEQ ID NO:5644.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;

XX colorectal carcinoma.

XX Homo sapiens.

XX WO200122920-A2.

XX 05-APR-2001.

XX 28-SEP-2000; 2000WO-US026524.

XX 29-SEP-1999; 99US-0157137P.

XX 03-NOV-1999; 99US-0163280P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Barash SC, Birse CE, Rosen CA;

XX WPI; 2001-235357/24.

XX N-PSDB; AAH34285.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,

XX useful for preventing, diagnosing and/or treating colorectal cancers.

XX Claim 11; Page 7194-7195; 9803pp; English.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon

XX cancer-associated nucleic acid molecules (N) and proteins (P), where the

XX proteins are collectively known as colon cancer antigens. The colon

XX cancer antigens have cytostatic activity and can be used in gene therapy

XX and vaccine production. N and P may be used in the prevention, diagnosis

XX and treatment of diseases associated with inappropriate P expression. For

XX example, N and P may be used to treat disorders associated with decreased

XX expression by rectifying mutations or deletions in a patient's genome

CC that affect the activity of P by expressing inactive proteins or to
CC supplement the patient's own production of P. Additionally, N may be used
CC to produce the colon cancer-associated Ps, by inserting the nucleic acids
CC into a host cell and culturing the cell to express the proteins. N and P
CC can be used in the prevention, diagnosis and treatment of colorectal
CC carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent
CC sequences used in the exemplification of the present invention. N.B.
CC Pages 666 to 682 and page 7053 of the sequence listing were missing at
CC time of publication, meaning no sequences are present for SEQ ID NO:1027
CC to 1052, 7921 and 7922
XX

SQ Sequence 407 AA;

Alignment Scores:

Pred. No.: 1,92e-194 Length: 407
Score: 1929.00 Matches: 364
Percent Similarity: 98.11% Conservatives: 0
Best Local Similarity: 98.11% Mismatches: 7
Query Match: 98.32% Indels: 0
DB: 4 Gaps: 0

US-10-046-935-2234 (1-1116) x AAG74880 (1-407)

QY 1 ATGGAGCGCTCGCGCGTGGCGAGAAAGATCTCAGAGTAAAGAAAGAACTTAAAGAAATTC 60
DB 37 MetAspAlaArgValProGlnLysAspLeuArgValLysAsnLeuLysLysPhe 56
QY 61 AGATATGTGAAGTTGATTTCCATGGAACCTGTCATCTCTCTGATGACAGTTGTGACAGC 120
DB 57 ArgTyrValLysLeuIleSerMetGluThrSerSerSerSerSerSerSerSerSerSer 76
QY 121 TTTGCTCTCTGATATTTTCGAAACACAGAGCTTCAGTTCGAGAGAGCTGTAGAGC 180
DB 77 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr 96
QY 181 CGCAGCAGTGCAGGCACTCTGGACCTCTCAGGCTGGCGATGAAGTTTCCAGCGCGAGT 240
DB 97 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 116
QY 241 ACCAGGGAGCAACCAACAAAAGCAGAGTCCCGCAGCCCTCAGAGAATTTCTGTGACT 300
DB 117 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 136
QY 301 GATTCCAACTCCGATTCAGAGATGAAGTGGATGAATTTTGGAGAAAAGGGCTTTA 360
DB 137 AppSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 156
QY 361 AATATAAGCAAAACAAAGCAATGCTTGCAAACTCATGTCTGAATTAGAAAGCTTCCCT 420
DB 157 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 176
QY 421 GGCTCGTTCGGTGAAGAGATCCCTCCAGGCTCCGACTCACAATCAAGAGAGACCGCGA 480
DB 177 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 196
QY 481 AGCGGTACATTCCTGGGTGCTTCCAGAGAAAACCTGACAGGAGAGCTGTCTCTCTT 540
DB 197 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgAlaArgProLeu 216
QY 541 ACCAGTCAAGTCCCGGATCCTCGGTCCTTGACGCTCTACCCATGCGAGAGCACT 600
DB 217 ThrArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMetGluGluGlu 236
QY 601 TTKGARKMTMWGTACATGTTGTGTGAGAAAGAGAGACCGTGGATGGCTACATGATGAA 660
DB 237 GluGluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 256
QY 661 GATGACCTGCCAGAGCGCTCGCTCCAGATCATCGTGACCTCCCGCATATTAATTCG 720
DB 257 AspAspLeuProArgSerArgArgSerArgSerValThrLeuProHisIleLeuArg 276
QY 721 CCAGTGGAGAAATTTACAGAGGAGCTTGGAGAGAGCTCTGAGCAATTTCTCGAGAGAAG 780

Db 277 ProValGluGluIleThrGluGluGluLeuGluAenValCysSerAsnSerArgGluLys 296
QY 781 ATATATAACCGTTCACTGGCTCTACTTGTCTATCAATCCGTCAGAACACTATTGATACC 840
Db 237 IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 316
QY 841 AAAACAAACTGCAGAACCCAGACTGCTGGGGCGTTTCGAGGCCAGTTCTGTGGCCCTGC 900
Db 317 LysThrAsnCysArgAsnProAspCysTrpGlyValArgGlyGlnPheCysGlyProCys 336
QY 901 CTTGAAACCGTTATGGTGAAGAGTTCAGGGATGCTCTGCTGGATCCGAACTGGCATTCG 960
Db 337 LeuArgAsnArgTyrGlyGluGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 356
QY 961 CGCCTTGTTCGAGGAATCTGCAACTGCAGTTTCTCGGCGACGAGATGAGCGTGTGG 1020
Db 357 ProProCysArgGlyCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 376
QY 1021 ACTGGGCTCTTGTGTATTTAGCCAAATATCATGGCTTTGGGAATGTGCATCGCTACTTG 1080
Db 377 ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 396
QY 1081 AAAAGCTGAACACAGAAATTCGAATGCAAGCA 1113
Db 397 LysSerLeuLysGlnGluPheGluMetGlnAla 407

RESULT 9

ADA98284

ID ADA98284 standard; protein; 397 AA.

XX AC

ADA98284;

XX DT

20-NOV-2003 (first entry)

XX DE

Human secreted protein sequence #125.

XX KW

human; secreted protein; cardiovascular disorder; arrhythmia;
atherosclerosis; stroke; endocarditis; congestive heart failure;
rheumatic heart disease; cardiomyopathy; haemorrhoids; varicose veins;
migraine; thrombosis; neural disorder; immune system disorder;
muscular disorder; reproductive disorder; gastrointestinal disorder;
pulmonary disorder; renal disorder; proliferative disorder; cancer.

XX OS

Homo sapiens.

XX XX

WQ2003004623-A2.

XX PD

16-JAN-2003.

XX XX

26-MAR-2002; 2002WO-US0009922.

XX PF

27-MAR-2001; 2001US-0278650P.

XX PR

12-SEP-2001; 2001US-00950082.

XX PR

12-SEP-2001; 2001US-00950083.

XX XX

(HUMA-) HUMAN GENOME SCI INC.

XX PA

Rosen CA, Ruben SM;

XX PI

WPI; 2003-247946/24.

XX DR

New human secreted polypeptide and nucleic acid molecules, useful for
diagnosing, preventing, prognosticating or treating cardiovascular
disorders (e.g. arrhythmia, atherosclerosis, cardiomyopathy, or
thrombosis).

XX PT

Claim 1; SEQ ID NO 392; 1572pp; English.

XX PS

The invention comprises the amino acid and coding sequence of human
secreted proteins. The DNA and protein sequences of the invention are
useful in the treatment of cardiovascular disorders, such as: arrhythmia,
atherosclerosis, stroke, endocarditis, congestive heart failure,
rheumatic heart disease, cardiomyopathy, haemorrhoids, varicose veins,

CC migraine, or thrombosis. The DNA and protein sequences may also be used
for treating or preventing: neural disorders, immune system disorders,
CC muscular disorders, reproductive disorders, gastrointestinal disorders,
CC pulmonary disorders, renal disorders, proliferative disorders and/or
CC cancerous diseases. The present amino acid sequence represents a human
secreted protein of the invention. NOTE: The present sequence is shown on
CC the WIFO website.

XX

SQ Sequence 397 AA;

Alignment Scores:

Pred. No.: 6,42e-194 Length: 397
Score: 1924.00 Matches: 364
Percent Similarity: 97.85% Conservative: 0
Best Local Similarity: 97.85% Mismatches: 7
Query Match: 98.06% Indels: 1
DB: Gaps: 0

US-10-046-935-2234 (1-1116) x ADA98284 (1-397)

QY 1 ATGGACGCTCGCGCGTGCAGAAAGATCTCAGAGTAAAGAACTTAAGAATTC 60

Db 1 MetAspAlaArgValProGlnLysAspLeuArgValLysLysAsnLeuLysLysPhe 20

QY 61 AGATATGTGAAGTTGATTTCCATGGAACCTCGTCTCTCTGATCAGAGTTGTGACAGC 120

Db 21 ArgTyrValLysLeuIleSerMetGluThrSerSerSerSerSerSerSerSerSer 40

QY 121 TTTTGTCTCTGATAATTTTGCACACAGAGCTGCAGTTCAGTTTCGGGAAGCTGTAGAGC 180

Db 41 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGlyCysArgThr 60

QY 181 CGCAGCAGTGCAGGCACCTCTCGACCTCTCAGGGTGGCGATGAAGTTTCAGCGCGAGT 240

Db 61 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 80

QY 241 ACAGGGAGCAACCAACAAAAGCAGAGTCCCGCCGCTCAGAGATTCGTGACT 300

Db 81 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 100

QY 301 GATTCACATCCGATTTCAGAGTGAAGTGAATGAATTTTGGAGAGAAAGGGCTTTA 360

Db 101 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 120

QY 361 AATATAAGCAAAACAAACCAATGTTTGCACAACTCATGTCTGAAATAGAAAGCTTCCCT 420

Db 121 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140

QY 421 GGCTCGTTCCGTGGAAGACATCCCTCCAGGCTCCGACTCACAATCAAGGAGCCGCGA 480

Db 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 160

QY 481 AGCGGTACATTCGCGGTGTTGCTTCAGAGAGAAACCTGAAACGAGAGCTCGTCTCTT 540

Db 161 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgArgAlaArgProLeu 180

QY 541 ACCAGTCAAGTCCCGGATTCCTCGGTCCTTTCAGCTTACCTTCCATGCGAGAGCCAGT 600

Db 181 ThrArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMetGluGluGlu 200

QY 601 TTKGARKMTWGTACATGTTGGTGAAGAGCAACCGTGGATGGCTACATGAATGAA 660

Db 201 GluGluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 220

QY 661 GATGACCTGCCAGAGCGCTCGCTCCAGATCATCGTGCACCTTCGCGATATAATTCGC 720

Db 221 AspAspLeuProArgSerArgArgSerArgSerValThrLeuProHisIleArg 240

QY 721 CCAGTGAAGAAATTCAGAGAGAGGAGTGGAGAAAGCTGTGACCAATTCGAGAGAG 780

Db 241 ProValGluGluIleThrGluGluGluLeuGluAsnValCysSerAsnSerArgGluLys 260

QY 781 ATATATAACCGTTCACTGGGCTCTACTTGTTCATCAATGCCGTGAGAGAGACTATTGATACC 840

Db 261 IIEYRASNARSGERLEUGLYSERTHRYSHISGLNCYSARGGLNLYSTHRIEASPTH 280
QY 841 AAAACAAACTGCAGAAACCCAGACTGCTGGGCGCTTCGAGGCGAGTCTCTGGCCCTGC 900
Db 281 LysThrAsnCysArgAsnProAspCysTrpGlyValArgGlyGlnPheCysGlyProCys 300
QY 901 CTTGCAAAACCGTTATGGTGAAGAGTCAGGGATGCTCTGCTGGATCCGAACCTGGCATTGC 960
Db 301 LeuArgAsnArgTyrGlyGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 320
QY 961 CGCCTCTGTCGAGGAATCTGCAACTGCAGTTCTCCGCGCAGCAGATGAGCGGTGTGCG 1020
Db 321 ProProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 340
QY 1021 ACTGGGTCCTTGTGTATTTAGCCAAATATCATGCTTTTGGAAATGTCATGCCCTACTTG 1080
Db 341 ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 360
QY 1081 AAAAGCCTGAACAGGAATTTGAATGCAAGCAT 1114
Db 361 LysSerLeu-AsnArgAsnLeuLysCysLysHis 371
RESULT 10
ADA44119
ID ADA44119 standard; protein; 397 AA.
XX
AC ADA44119;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human secreted protein SEQ ID 311.
XX
KW Gene therapy; human; Antidiabetic; Anorectic; Ophthalmological;
KW Neuroprotective; Cerebroprotective; Antianemic.
XX
OS Homo sapiens.
XX
PN W0203000865-A2.
XX
PD 03-JAN-2003.
XX
PF 26-MAR-2002; 2002WO-US009105.
XX
PR 27-MAR-2001; 2001US-0278650P.
PR 12-SEP-2001; 2001US-00950082.
PR 12-SEP-2001; 2001US-00950083.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2003-184045/18.
DR N-PSDB; ADA43929.
XX
PT A human secreted protein and nucleic acids useful for preparing a
PT diagnostic or pharmaceutical composition for diagnosing or treating
PT diabetes or conditions related to diabetes, e.g. hyperglycemia, obesity,
PT retinopathy, neuropathy.
XX
PS Claim 1; SEQ ID NO 311; 701pp; English.
XX
CC The invention relates to novel genes and their fragments which are useful
CC for preventing, treating or ameliorating medical conditions e.g. by
CC protein or gene therapy. The genes are isolated from a range of human
CC tissues disclosed in the specification. The nucleic acids and proteins
CC are useful in the diagnosis, treatment and prevention of conditions
CC related to diabetes, e.g. hyperglycaemia, obesity, retinopathy,
CC polynuropathy, atherosclerosis, anaemia, stroke, gangrene, impotence,
CC infection, cataract, renal disorders, or endocrine disorders. The present
CC sequence was used to illustrate the invention.
XX
SQ Sequence 397 AA;

Alignment Scores:
Pred. No.: 6,42e-194 Length: 397
Score: 1924.00 Matches: 364
Percent Similarity: 97.85% Conservative: 0
Best Local Similarity: 97.85% Mismatches: 7
Query Match: 98.06% Indels: 1
DB: 6 Gaps: 0
US-10-046-935-2234 (1-1116) x ADA44119 (1-397)
QY 1 ATGACGCTCCGCGCTGCCGAGAAAGATCTCAGAGTAAGAAGAACTTAAGAAATTC 60
Db 1 MetAspAlaArgValProGlnLysAspLeuArgValLysLysAsnLeuLysPhe 20
QY 61 AGATATGGAAGTTGATTTCATCGAAACCTCGTCATCTCTGATGACAGTTGTGACAGC 120
Db 21 ArgTyrValLysLeuIleSerMetGluThrSerSerSerSerSerSerSerSerSer 40
QY 121 TTTGCTTCTGATTAATTTTGCACACGAGGCTGCAGTCAGTTCGGGAAAGCGTGTAGGACC 180
Db 41 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr 60
QY 181 CGCAGCCAGTCAGGCACTCTGGACCTCTCAGGCTGGCGATGAGTATTCACCGCGGAGT 240
Db 61 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 80
QY 241 ACAGGGGAGCAACAAACAAAGAGAGTCCCGCCAGCCCTCAGAGAAATTCGTGACT 300
Db 81 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 100
QY 301 GATTCCAACTCCGATTTCAGAGATGAAGTGAATGAAATTTTGGAGAAAAGCGCTTTA 360
Db 101 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 120
QY 361 AATATAAGCAAAACAAAGCAATGCTTGCAAACTCATGTCTGAATTAGAAGCTTCCT 420
Db 121 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140
QY 421 GGCTGCTGCTGGAGACATCCCTCCAGGCTCCGACTCACAATCAAGAGAGACCCGA 480
Db 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 160
QY 481 AGCGTACATTCGCGGTGTGCTTCAGAGAGAAACCTCGAAGAGAGTGTGCTCTT 540
Db 161 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgArgAlaArgProLeu 180
QY 541 ACCAGGTCAAGTCCCGGATCCTCGGCTCCCTGAGCTCTACCCATGGRGAGGACAGT 600
Db 181 ThrArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMetGluGluGlu 200
QY 601 TTGARKMTWGTACATGTTGGTGAAGAGAGACCGTGGATGGCTACATGAATGAA 660
Db 201 GluGluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 220
QY 661 GATGACCTGCCAGAGCCGTCCTCCAGATCATCGTGACCTTCGCGATATAATTCGC 720
Db 221 AspAspLeuProArgSerArgArgSerArgSerValThrLeuProHisIleIleArg 240
QY 721 CCAGTGGAAAGAAATTAACAGAGGAGGAGTGGAGAGCTGTGCAGCAATTCCTCAGAGNAG 780
Db 241 ProValGluGluIleThrGluGluGluLeuGluAsnValCysSerAsnSerArgGluLys 260
QY 781 ATATATAACCGTTCACCTGGGCTCTACTTGTCAATGCCGTCCAGAGAACTATTGATACC 840
Db 261 IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 280
QY 841 AAAACAACTGCAGAAACCCAGCTGCTGGGCGGTTCGAGGCGAGTCTGTGCGCCCTGC 900
Db 281 LysThrAsnCysArgAsnProAspCysTrpGlyValArgGlyGlnPheCysGlyProCys 300
QY 901 CTTGCAAAACCGTTATGGTGAAGAGGTCAGGAGTGTCTGCTGGATCCGAGTTCG 960

Db 301 LeuArgAsnArgTyrGlyGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 320
 QY 961 CGCCTTGTGAGGAAATCTGCACTGCTTCTCGCGCAGCAGATGCGGTGCG 1020
 Db 321 ProProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 340
 QY 1021 ACTGGGCTTGTGTAATATGCAATATCATGCTTTGGGAATGTGATCGCTACTTG 1080
 Db 341 ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 360
 QY 1081 AAAAGCTGAAACAGGAATTTGAAATGCAAGCAT 1114
 Db 361 LysSerLeu-AsnArgAsnLeuLysCysLysHis 371
 RESULT 11
 ADC20449
 ID ADC20449 standard; protein; 397 AA.
 XX
 AC ADC20449;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE Human secreted protein - amino acid sequence #130.
 XX
 KW gene therapy; human; secreted protein; haemopoietic disorder;
 KW haematological disorder; anaemia; haemophilia; inflammatory disorder;
 KW inflammatory bowel disease; Crohn's disease; neoplastic disease; cancer;
 KW leukemia; wound healing; epithelial cell proliferation disorder;
 KW immune disorder; autoimmune disorder; asthmatic disorder;
 KW cardiovascular disorder; atherosclerosis; myocarditis;
 KW infectious disease; HIV; AIDS; endocrine disorder; diabetes;
 KW gastrointestinal disorder; duodenal ulcer; gastroenteritis.
 XX
 OS Homo sapiens.
 XX
 PN WC020292787-A2.
 XX
 PD 21-NOV-2002.
 XX
 PF 26-MAR-2002; 2002WO-US009257.
 XX
 PR 27-MAR-2001; 2001US-0278650P.
 PR 12-SEP-2001; 2001US-00950082.
 PR 12-SEP-2001; 2001US-00950083.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI; 2003-123287/12.
 XX
 PT New human secreted proteins and nucleic acid molecules, useful for
 PT preparing a diagnostic or pharmaceutical composition for diagnosing,
 PT preventing or treating hematopoietic or hematologic disorders, e.g.
 PT anemia or hemophilia.
 XX
 FS Claim 1; SEQ ID NO 403; 1512pp; English.
 XX
 CC The invention comprises the amino acid and coding sequences of human
 CC secreted proteins. The DNA and protein sequences of the invention are
 CC useful for detecting, preventing, diagnosing, prognosticating, treating
 CC or ameliorating: haematopoietic or haematological disorders (e.g. anaemia
 CC and haemophilia); inflammatory disorders (e.g. inflammatory bowel disease
 CC and Crohn's disease); neoplastic disease (e.g. cancer and leukaemia);
 CC wound healing and disorders of epithelial cell proliferation; immune
 CC disorders (e.g. autoimmune disorders and asthmatic disorders);
 CC cardiovascular disorders (e.g. atherosclerosis and myocarditis);
 CC infectious disease (e.g. HIV/AIDS); endocrine disorders (e.g. diabetes);
 CC and gastrointestinal disorders (e.g. duodenal ulcers and
 CC gastroenteritis). The present amino acid sequence represents a human
 CC secreted protein of the invention.
 XX
 SQ Sequence 397 AA;

Alignment Scores:
 Pred. No.: 6,42e-194 Length: 397
 Score: 1924.00 Matches: 364
 Percent Similarity: 97.85% Conservative: 0
 Best Local Similarity: 97.85% Mismatches: 7
 Query Match: 98.06% Indels: 1
 DB: 7 Gaps: 0
 US-10-046-935-2234 (1-1116) x ADC20449 (1-397)
 QY 1 ATGAGCGCTCGCGGTGCGCGCAGAAAGATCTCAGATGAAAGAACTTAAAGAATTC 60
 Db 1 MetAspAlaArgArgValProGlnLysAspLeuArgValLysLysLysLysPhe 20
 QY 61 AGATATGTGAAGTTGATTTCCATGAAACCTCGTCACTCTGATGACAGTTGTGACGC 120
 Db 21 ArgTyrValLysLeuIleSerMetGluThrSerSerSerSerAspAspSerCysAspSer 40
 QY 121 TTTGCTTCTGATAATTTTGCACACAGAGCGTGCAGTTCGGGAAAGCTGTAGGACC 180
 Db 41 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr 60
 QY 181 CGCAGCGAGTGCAGGCACTCTGCACCTCTCAGGCTGCGCATGAAGTTTCCAGCGCGGAGT 240
 Db 61 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 80
 QY 241 ACCAGGGAGAGCAACCAACAAAGAGAGTCCCGCCAGCCCTCAGAGAAATTTCTGACT 300
 Db 81 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 100
 QY 301 GATTCCAACTCCGATTCAGAGATGAAGTGAATGATTTTGGGAGAAAGGGCTTTA 360
 Db 101 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 120
 QY 361 AATATAAGCAAAACAAAGCAATGCTTGCAAACTCATGTCTCAATTAGAAAGCTTCCT 420
 Db 121 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140
 QY 421 GGCTCGTTCGTTGGAGACATCCCTCCAGGCTCCGACTCACAATCAAGAGAGACCGGA 480
 Db 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 160
 QY 481 AGCGTACATTCGGGTGTGTTCTCCAGAGAAACCTGAACGGAGAGCTCGTCTCTT 540
 Db 161 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgArgAlaArgProLeu 180
 QY 541 ACCAGGTCAGGTCCTCGGATCCCTCGGCTCCCTTACCGCTCTACCCATGGRGAGAGCCAGT 600
 Db 181 ThrArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMetGluGluGlu 200
 QY 601 TTKGARKMTWGTACATGTTGGTGAGAAAGAGACCGGTGATGCTACATGAATGAA 660
 Db 201 GluLysAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 220
 QY 661 GATGACCTGCCAGAGCGCTCGCTCCAGATCATCGTGCACCTTCGSCATATATTCGC 720
 Db 221 AspAspLeuProArgSerArgSerValThrLeuProHisIleLeuArg 240
 QY 721 CCAGTGGAGAAATTTACAGAGAGGAGTTCGAGAACGTTCTGCAGCAATTTCTCGAGAGAG 780
 Db 241 ProValGluGluIleThrGluGluGluLeuGluAsnValCysSerAsnSerArgGluLys 260
 QY 781 ATATATAACCGTTCACTGGGTCTACTTGTGTCATCATCGCTCAGAGACACTATTGATACC 840
 Db 261 IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 280
 QY 841 AAAACAACCTCAGAAACCCAGACTCTGGGGCGTTCGAGGCGAGTTCGTGGCCCGCTGC 900
 Db 281 LysThrAsnCysArgAsnProAspCysTrpGlyValArgGlyGlnPheCysGlyProCys 300
 QY 901 CTTTCGAAAACCGTTATGGTGAAGAGTTCAGGGATGCTCTGCTGGATCCGAACTGGCATTCG 960

Dd 301 LeuArgAsnArgTyrGlyGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 320
 Qy 961 CCGCTTGTGAGGAATCTGCAACTGAGTTTCTCCGGCAGCAGAGATGGAGTGTGCG 1020
 Dd 321 ProProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysala 340
 Qy 1021 ACTGGGCTCTTGTGTTATTTAGCCAAATATCATGGCTTTGGGAATGTGCATGCTACTTG 1080
 Dd 341 ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 360
 Qy 1081 AAAAGCCTGAAACAGGAAATTTGAATGCAAGCAT 1114
 Dd 361 LysSerLeu-AsnArgAsnLeuLysCysLysHis 371

RESULT 12

ADF10748

ID ADF10748 standard; protein; 397 AA.

XX AC ADF10748;

XX AC ADF10748;

XX AC ADF10748;

DT 12-FEB-2004 (first entry)

XX Human secreted protein #70.

XX H6EDM64; HBHAA05; HBJCR46; HBJKD16; HCMXS51; HCQBH72; HDPPQ30; HE2CM39;

XX H29EA10; HGBHP91; HLDQU79; Cytostatic; Hepatotropic; Antidiabetic;

XX Antiinflammatory; neuroprotective; Anti-HIV; Vulnerary; Gynecological;

XX Antinfertility; Gene therapy; gastrointestinal disorder; cancer;

XX Alzheimer's disease; chromosome identification.

XX Homo sapiens.

XX OS Homo sapiens.

XX PN WO200299085-A2.

XX 12-DEC-2002.

XX 26-MAR-2002; 2002WO-US0009135.

XX 27-MAR-2001; 2001US-0278650P.

XX 12-SEP-2001; 2001US-00950082.

XX 12-SEP-2001; 2001US-00950083.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2003-221310/21.

XX New human secreted polypeptides for diagnosing and treating neural,

XX immune system, muscular, reproductive, gastrointestinal, cardiovascular,

XX renal, and proliferative disorders and cancerous diseases.

XX Claim 1; SEQ ID NO 211; 855pp; English.

XX The present invention relates to an isolated polypeptide chosen from 123

XX human secreted proteins, such as, H6EDM64, HBHAA05, HBJCR46, HBJKD16,

XX HCMXS51, HCQBH72, HDPPQ30, HE2CM39, H29EA10, HGBHP91 and HLDQU79. The

XX polypeptides are useful for the preparation of a diagnostic or

XX pharmaceutical composition for diagnosing or and are useful for treating

XX or preventing diseases or conditions, such as neural, immune system,

XX muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,

XX renal, proliferative disorders and cancerous diseases and conditions. The

XX polypeptides have immune activity, chemotactic activity, and binding

XX activity to treat and prevent neuronal damage which occurs in certain

XX neuronal disorders or neuro-degenerative conditions such as Alzheimer's

XX disease, Parkinson's disease, and acquired immunodeficiency syndrome

XX (AIDS)-related complex, and to prevent skin aging due to sunburn by

XX stimulating keratinocyte growth. The molecules are also useful to

XX modulate mammalian characteristics including .The encoding sequences are

XX useful for chromosome identification, radiation hybrid mapping, in gene

XX therapy, for identifying individuals from minute biological samples, as

XX additional DNA markers for restriction fragment length polymorphism

XX (RFLP), in forensic biology, molecular weight markers on Southern gels,

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

CC as diagnostic probes for the presence of a specific mRNA in a particular
 CC cell type, to raise anti-DNA antibodies using DNA immunization
 CC techniques, and as an antigen to elicit an immune response. The present
 CC sequence represents a human secreted protein of the invention.
 XX

SQ Sequence 397 AA;

Alignment Scores:

Pred. No.: 6,42e-194 Length: 397
 Score: 1924.00 Matches: 364
 Percent Similarity: 97.85% Conservative: 0
 Best Local Similarity: 97.85% Mismatches: 7
 Query Match: 98.06% Indels: 1
 DB: 7 Gaps: 0

US-10-046-935-2234 (1-1116) x ADF10748 (1-397)

Qy 1 ATGACGCTCCGCGCGTGGCCAGAAAGATCTCAGAGTAAGAAGAACTTAAAGAAATTC 60
 Dd 1 MetAspAlaArgValProGlnLysAspLeuArgValLysLysAsnLeuLysLysPhe 20
 Qy 61 AGATATGTGAAGTGTATTTCCATGGAAACCTCGTCATCTCTGATGACAGTTGTGACAGC 120
 Dd 21 ArgTyrValLysLeuIleSerMetGluThrSerSerSerSerSerSerSerSerSerSer 40
 Qy 121 TTTGCTTCTGATTAATTTTGCACACAGAGCTCGAGTCAGTTCGGGAAGGCTGTAGGACC 180
 Dd 41 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr 60
 Qy 181 CGCAGCAGTGCAGGCACTCTGACCTCTCAGGCTGGCGATGAAGTTTCAGAGCCGAGT 240
 Dd 61 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 80
 Qy 241 ACCAGGGAGCAACCAACAAAAAGCAGAGTCCCGCCAGCCCTCAGAGAAATTTGTGACT 300
 Dd 81 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 100
 Qy 301 GATTCCAACCTCCGATTCAGAGAGTGAAGTGAATGAATTTTGGAGAAAAGGGCTTTA 360
 Dd 101 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 120
 Qy 361 AATATAAGCAAAACAAAGCAATGCTTGCAAAACCTCATGTCTGAATTAGAAAGCTTCCCT 420
 Dd 121 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140
 Qy 421 GGCTCGTTCCTGGGAAGACATCCCTCCAGGCTCCGACTCACAATCAAGGAGACCGCGA 480
 Dd 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 160
 Qy 481 AGCGGTACATTCGCGGTCTTCTCCAGGAGAAACCTTGAACGGAGAGCTGCTCTCTT 540
 Dd 161 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgAlaArgProLeu 180
 Qy 541 ACCAGTCAAGTCCCGATCTCCGGTCCCTTGCAGCTCTACCCATGGRGAGAGCCACT 600
 Dd 181 ThrArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuPrometGluGluGlu 200
 Qy 601 TTKGARKMTWGTACATGTTGTGAGAAGAGAGACCGTGGATGGCTTACATGAATGAA 660
 Dd 201 GluGluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 220
 Qy 661 GATGACCTGCCAGAGCGGTCCGTCAGATCATCCGTGACCTTCCGATATATTCG 720
 Dd 221 AspAspLeuProArgSerArgArgSerValThrLeuProHisIleArg 240
 Qy 721 CCAGTGGGAAGAAATACAGAGAGAGTGGAGAGAGCTGCAGCAATTCGACGAGAG 780
 Dd 241 ProValGluGluIleThrGluGluGluLeuAsnValCysSerAsnSerArgGluLys 260
 Qy 781 ATATATAACCGTTCACCTGGGTCTTACTGTTCATCAATGCCGTTCAGAGCAATTCATACC 840
 Dd 261 IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 280

QY 841 AATCAAACTGAGAAACCCAGACTGCTGGGCGTTCGAGGCCAGTCTCTGGCCCTTCG 900
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 281 LysThrAsnCysArgAsnProAspCysTrpGlyValAlaGlyGlnPheCysGlyProCys 300
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 901 CTTGAAACCCGTTATGCTGAAGAGCTCAGGAGTCTCTCTGGATCCGAATCGCATTCG 960
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 301 LeuArgAsnArgTyGlyGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 320
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 961 CGCTCTGTGAGGAATCTGCACTGCACTTCTCCCGCAGCAGATCGAGCGTGGG 1020
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 321 ProProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 340
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1021 ACTGGGCTCTGTGTTATTTAGCCAAATATCATGCTTTGGGAATGTGATCCCTACTTG 1080
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 341 ThrGlyValLeuValTyrrLeuAlaIleTyrrHisGlyPheGlyAsnValHisAlaTyrrLeu 360
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1081 AAAAGCCTGAAACAGCAATTTGAAATGCAAGCAT 1114
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 361 LysSerLeu-AsnArgAsnLeuLysCysLysHis 371
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
RESULT 13
ID AAB93122 standard; protein; 450 AA.
XX AAB93122;
XX 26-JUN-2001 (first entry)
XX Human protein sequence SEQ ID NO:12001.
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX Homo sapiens.
XX EF1074617-A2.
XX 07-FEB-2001.
XX 28-JUL-2000; 2000EP-00116126.
XX 29-JUL-1999; 99JP-00248036.
XX 27-AUG-1999; 99JP-00300253.
XX 11-JAN-2000; 2000JP-00118776.
XX 02-MAY-2000; 2000JP-00183767.
XX 03-JUN-2000; 2000JP-00241899.
XX (HELI-) HELIX RES INST.
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
XX Claim 8; SEQ ID NO 12001; 2537pp + Sequence Listing; English.
XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises:
XX (a) an oligo-dT primer and an oligonucleotide complementary to the
XX complementary strand of a polynucleotide which comprises one of the 5602
XX nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in the
XX specification. The primer sets can be used in antisense therapy and in

CC gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the CC detection and/or diagnosis of the abnormality of the proteins encoded by CC the full-length cDNAs. The primers allow obtaining of the full-length CC cDNAs easily without any specialised methods. AAB93122 to AAB95893 CC AAB93122 to AAB95893 represent human cDNA sequences; AAB92446 to AAB95893 CC represent human amino acid sequences; and AAB9329 to AAB93632 represent CC oligonucleotides, all of which are used in the exemplification of the CC present invention
XX
XX Sequence 450 AA;
SQ
Alignment Scores:
Pred. No.: 3,498-189 Length: 450
Score: 1879.50 Matches: 364
Percent Similarity: 80.8% Conservatives: 0
Best Local Similarity: 80.8% Mismatches: 7
Query Match: 95.8% Indels: 79
DB: 4 Gaps: 1
US-10-046-935-2234 (1-1116) x AAB93122 (1-450)
QY 1 ATGGAGCTCGCGCTGCGCGAGAAAGATCTCAGAGTAAAGAACTTAAAGAAATTC 60
DB 1 MetAspAlaArgValProGlnLysAspLeuArgValLysLysAsnLeuLysLysPhe 20
QY 61 AGATATGTGAAGTTGATTTCCATGGAACCTCGTCATCTCTGATGACAGTTGTGACAGC 120
DB 21 ArgTyrrValLysLeuIleSerMetGluThrSerSerSerSerSerSerSerSerSer 40
QY 121 TTTGCTTCTGTAATTTTTCRAACAGC----- 147
DB 41 PheAlaSerAspAsnPheAlaAsnThrLysProLysPheArgSerAspLeuSerGluGlu 60
QY 147 ----- 147
DB 61 LeuAlaSerValPheTyrrGluAspSerAspAsnGluSerPheCysGlyPheSerGluSer 80
QY 147 ----- 147
DB 81 GluValGlnAspValLeuAspHisCysGlyPheLeuGlnLysProArgProAspValThr 100
QY 147 ----- 147
DB 101 AsnGluLeuAlaGlyIlePheHisAlaAspSerAspAspGluSerPheCysGlyPheSer 120
QY 148 -----AGGCTGCAGTTCAGTTCGGAAGGCTGTAGACCCGC 183
DB 121 GluSerGluLeuGlnAspGlyMetArgLeuGlnSerValArgGluGlyCysArgThrArg 140
QY 184 AGCCAGTGCAGGCACCTCTGACCTCTCAGGTTGGCGATGAAGTTTCAGGCGCGAGTACC 243
DB 141 SerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSerThr 160
QY 244 AGGGAGCACCAACAAAGCAATGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 303
DB 161 ArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThrAsp 180
QY 304 TCCAACTCCGATTCAGAAAGATGAAAGTGGATGAAATTTTGGAGAAAAGGCTTTAAAT 363
DB 181 SerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeuAsn 200
QY 364 ATAAGCAAAACAAAGCAATGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 423
DB 201 IleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPheProGly 220
QY 424 TCGTTCCGTGGAGACATCCCTCCAGGCTCCGACTCACAATCAAGAGACCCGAGG 483
DB 221 SerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArgArg 240
QY 484 CGTACATTCGCGGTGTGTGCTTCAGGAGAAACCTCAAGAGAGCTGCTGCTCTTACC 543
DB 241 ArgThrPheProGlyValAlaSerArgArgAsnProGluArgAlaArgProLeuThr 260

QY 1 ATGGACGCTCGCGTCCGAGAAAGATCTCAGAGTAAAGAACTTAAAGAAATTC 60
 Db 1 MetAspAlaIaArgValProGlnuysAspLeuAArgValLysAsnLeuLysPhe 20
 QY 61 AGATATGTGAAGTTGATTTCCATGAAACCTCGTCACTCTGTGATGACAGTGTGACAGC 120
 Db 21 ArgTyrValLysLeuLysMetGluThrSerSerSerSerSerSerSerSerSerSer 40
 QY 121 TTTGCTTCTGATAAATTTTGCACACG 147
 Db 41 PheAlaSerAspAsnPheAlaAsnThrLysProLysPheArgSerAspLysSerGluGlu 60
 QY 147 147
 Db 61 LeuAlaSerValPheTyrGluAspSerAspAsnGluSerPheCysGlyPheSerGluSer 80
 QY 147 147
 Db 81 GluValGlnAspValLeuAspHisCysGlyPheLeuGlnLysProArgProAspValThr 100
 QY 147 147
 Db 101 AsnGluLeuAlaGlyPheHisAlaAspSerAspAspGluSerPheCysGlyPheSer 120
 QY 148 148
 Db 121 GluSerGluLeuGlnAspGlyMetArgLeuGlnSerValArgGluGlyCysArgThrArg 140
 QY 184 AGCAGTGCAGGCACTCTCGACCTCTCAGGTTGGCGATGAAGTTTCCAGCGGAGTACC 243
 Db 141 SerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSerThr 160
 QY 244 AGGGAGCAACCAACAAAGACAGATCCCGCAGCCCTCAGAGATTCGTGACTGAT 303
 Db 161 ArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThrAsp 180
 QY 304 TCCAACTCCGATTCAGAGTGAAGTGAATGAATTTTGGAGAAAGGGCTTTAAAT 363
 Db 181 SerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeuAsn 200
 QY 364 ATAAAGCAACAAAGCAATGCTTGCAAACTCATGTCTGAATAGAAAGCTTCCCTGCC 423
 Db 201 IleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPheProGly 220
 QY 424 TCGTTCCGTGGAAGACATCCCTCCAGGCTCCGACTCACAATCAAGGACCGCGAAGG 483
 Db 221 SerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArgArg 240
 QY 484 CGTACATTCGCGGTGTGCTCCAGGAGAAACCTGAAACGAGAGCTCGTCTCTTACC 543
 Db 241 ArgThrPheProGlyValAlaSerArgArgAsnProGluArgArgAlaArgProLeuThr 260
 QY 544 AGTCAAGTCCCGGATTCCTCGGTCCTTCCAGGCTTCCAGGCTTCCAGGAGCCAGTTK 603
 Db 261 ArgSerArgSerArgLysLeuLysSerLeuAspAlaLeuProMetGluGluGluGlu 280
 QY 604 GARKMTWGTACATGTTGGTGAAGAGAGAGACCGTGGATGCTACATGAATGAAGAT 663
 Db 281 GluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGluAsp 300
 QY 664 GACCTGCCAGAGACCCGCTCGCTCAGATCATCCGTGACCTTCGCGATATAATTCGCCCA 723
 Db 301 AspLeuProArgSerArgArgSerArgSerValThrLeuProHisLleArgPro 320
 QY 724 GTGGAAGAAATTCAGAGGAGGAGTTCGAGACGCTCGACCAATTCGAGAGAGATA 783
 Db 321 ValGluGluLysThrGluGluGluLeuGluAsnValCysSerAsnSerArgGluLysLe 340
 QY 784 TATAACGTTACCTGGGCTCTACTTGTTCATCAATCCGCTCAGAGAGATTTGATACCAA 843
 Db 341 TyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrLleAspThrLys 360

QY 844 ACAAACTCGAGAAACCCAGAGCTGCTGGGGCGTTTCAGGCCAGTTCTGTGGCCCTGCTT 903
 Db 361 ThrAsnCysArgAsnProAspCysTrpGlyValArgGlyGlnPheCysGlyProCysLeu 380
 QY 904 CGAAACCGTTATGGTGAAGAGCTCAGGATGCTCTGTGTGATCCGAACCTGGCATTTGCCCG 963
 Db 381 ArgAsnArgTyrGlyGluGluValArgAspAlaLeuLeuAspProAsnTrpHisCysPro 400
 QY 964 CTTTCTCGAGGAATCTGCAACTGCAGTTTCTGCCGCGAGAGATGGACGGTGTGCGACT 1023
 Db 401 ProCysArgGlyLysCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAlaThr 420
 QY 1024 GGGTCTCTGTGATTTAGCCAAATATCATGCGCTTTGGAAATGTCATGCTACTTGAAA 1083
 Db 421 GlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeuLys 440
 QY 1084 AGCCTGAAACAGGAATTTGAAATGCAAGCA 1113
 Db 441 SerLeuLysGlnGluPheGluMetGlnAla 450

RESULT 15

ADN06070

ID ADN06070 standard; protein; 450 AA.

AC ADN06070;

DT 01-JUL-2004 (first entry)

DE Antipsoriatic protein sequence #1189.

XX antipsoriatic; gene therapy; psoriasis; diagnosis.

XX Homo sapiens.

XX WO2004028479-A2.

XX 08-APR-2004.

XX 25-SEP-2003; 2003WO-US030907.

XX 25-SEP-2002; 2002US-0414006P.

XX (GETH) GENENTECH INC.

XX Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood W;

XX Wu TD;

XX WPI; 2004-305105/28.

XX N-PSDB; ADN06069.

XX New PRO nucleic acid or polypeptide, useful for preparing a

XX pharmaceutical composition for diagnosing or treating psoriasis in a

XX mammal.

XX Claim 9; SEQ ID NO 2465; 3069pp; English.

XX The invention relates to novel polynucleotide and polypeptides for

XX treating psoriasis or a sequence having at least 80% identity to the

XX above sequences. The nucleic acid is useful for preparing a composition

XX for diagnosing or treating psoriasis in a mammal. This sequence

XX corresponds to one of the polypeptides of the invention.

XX SQ Sequence 450 AA;

Alignment Scores:

Pred. No.: 3,49e-189 Length: 450

Score: 1879.50 Matches: 364

Percent Similarity: 80.89% Conservative: 0

Best Local Similarity: 80.89% Mismatches: 7

Query Match: 95.80% Indels: 79

DB: 8 Gaps: 1

US-10-046-935-2234 (1-1116) x ADN06070 (1-450)

```

QY 1 ATGACGCTCCGCGTCCGAGAAAGATCTCAGAGTAAAGAACTTAAAGAAATTC 60
Db 1 MetAspAlaArgArgValProGlnLysAspLeuArgValLysLysAsnLeuLysPhe 20
QY 61 AGATATGTGAATTCATTCATCGAAACCTCGTCATCCTCTGATGACAGTTGTGACAGC 120
Db 21 ArgTyrValLysLeuLysMetGluThrSerSerSerSerSerSerSerSerSerSer 40
QY 121 TTTGCTTCTGATAATTTTGCACACG----- 147
Db 41 PheAlaSerAspPheAlaAsnThrLysProLysPheArgSerAspLysSerGluGlu 60
QY 147 ----- 147
Db 61 LeuAlaSerValPheTyrGluAspSerAspAsnGluSerPheCysGlyPheSerGluSer 80
QY 147 ----- 147
Db 81 GluValGlnAspValLeuAspHisCysGlyPheLeuGlnLysProArgProAspValThr 100
QY 147 ----- 147
Db 101 AsnGluLeuAlaGlyIlePheHisAlaAspSerAspAspGluSerPheCysGlyPheSer 120
QY 148 -----AGCTGCAGTCAGTTCGGGAAGGCTGTAGGACCCGC 183
Db 121 GluSerGluIleGlnAspGlyMetArgLeuGlnSerValArgGluGlyCysArgThrArg 140
QY 184 AGCCAGTGCAGGCACCTCTGACACCTCTCAGGTGGCGATGAAGTTTCAGCGCGAGTACC 243
Db 141 SerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSerThr 160
QY 244 AGGGAGACCAACCAAAAGCAGAGTCCCGCAGCCCTCAGAGATTCCTGTGACTGAT 303
Db 161 ArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThrAsp 180
QY 304 TCCAATCCGATTTCAGAAATGAAGTGAATGAATTTTGGAGAAAAGGGCTTTAAAT 363
Db 181 SerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeuAsn 200
QY 364 ATAAAGCAAAACCAAGCAATGCTTGCAAACTCATGTCTGAATTAGAACTTCCCTGGC 423
Db 201 IleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPheProGly 220
QY 424 TCGTTCCGTGGAGACATCCCTCCAGGCTCCGACTCCGACTCACAATCAAGGACCGCGAAG 483
Db 221 SerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArgArg 240
QY 484 CGTACATTCGCGGTGTGCTCCAGGAGAAACCTGAAACGAGAGCTCGTCTCTTACC 543
Db 241 ArgThrPheProGlyValAlaSerArgArgAsnProGluArgArgAlaArgProLeuThr 260
QY 544 AGGTCAAGTCCCGGATCCTCGGCTCCCTGACGCTCTACCCATGCGAGAGCCAGTTTK 603
Db 261 ArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMetGluGluGluGlu 280
QY 604 GARKMTWGTACATGTGTGAGAAAGAGACCGTGGATGGCTACATGAATGAAGAT 663
Db 281 GluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGluAsp 300
QY 664 GACCTCCAGAAAGCGTCCGATCATCCGTGACCCCTCCGATATATAATTCGCCCA 723
Db 301 AspLeuProArgSerArgSerArgSerValThrLeuProHisIleIleArgPro 320
QY 724 GTGGAGAAATACAGAGGAGGTGGAGACCTCTGAGCAATTCCTGAGAGAGATA 783
Db 321 ValGluGluIleThrGluGluLeuGluAsnValCysSerAsnSerArgGluLysIle 340
QY 784 TATAACCGTTCACTGGGCTCTACTCTCATCAATCCGTCAGAGACTATTGATACCAA 843
Db 341 TyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThrLys 360

```

```

QY 844 ACRAACTGCAGAACCCAGACTGCTGGGCGCTTCGAGCCAGTTCTGTGGCCCTTCCTT 903
Db 361 ThrAsnCysArgAsnProAspCysTyrGlyValArgGlyGlnPheCysGlyProCysLeu 380
QY 904 CGAAACCGTTATGTGTAAGAGGTCAAGGATGCTCTGCTGGATCCGAACTGCATTGCCCG 963
Db 381 ArgAsnArgTyrGlyGluValArgAspAlaLeuLeuAspProAsnThrPheCysPro 400
QY 964 CTTTGTTCAGGAATCTGAACTGCACTGCTGCGGAGCGAGATGACGGTGTGCGACT 1023
Db 401 ProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAlaThr 420
QY 1024 GGGTCTCTGTGATTTAGCCAAATATCATGCTTTGGGAATGTGCATGCCCTACTTGA 1083
Db 421 GlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeuLys 440
QY 1084 AGCTGAAACAGGAATTTGAAATGCAAGCA 1113
Db 441 SerLeuLysGlnGluPheGluMetGlnAla 450

```

Search completed: October 12, 2004, 15:12:46
Job time : 127 secs

Blank

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 12, 2004, 15:18:46 ; Search time 100 Seconds
(without alignments)
7202.284 Million cell updates/sec

Title: US-10-046-935-2234

Perfect score: 1962
Sequence: 1 atgacgcgtgcgcgtgcc.....aattgaatgaagcataa 1116

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1356558 seqs, 322682953 residues

Total number of hits satisfying chosen parameters: 2713116

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool.p/US10046935/runat_12102004_155228_15327/app_query.fasta_1.1287
-DB=Published Applications AA -QFMT=fastan -SUFFIX=rapb -MINWATCH=0.1
-LOPCCI=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS-human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTPMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10046935 @CGN 1.153 @runat_12102004_155228_15327
-NCFU=6 -ICPU=3 -NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pcp.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pcp.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pcp.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pcp.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pcp.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pcp.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pcp.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pcp.*
18: /cgn2_6/ptodata/2/pubpaa/US10F_NEW_PUB.pcp.*
19: /cgn2_6/ptodata/2/pubpaa/US10G_NEW_PUB.pcp.*
20: /cgn2_6/ptodata/2/pubpaa/US10H_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	ID	Description
------------	-------	--------------	----	-------------

ALIGNMENTS

RESULT 1

US-09-878-178-2235
; Sequence 2235, Application US/09878178
; Patent No. US20020177552A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secretist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121 527
; CURRENT APPLICATION NUMBER: US/09/878,178
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 2237
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2235
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: (1)...(371)
; OTHER INFORMATION: Xaa = Any amino acid

1	1943	99.0	371	9	US-09-878-178-2235	Sequence 2235, Ap
2	1943	99.0	371	13	US-10-046-935-2235	Sequence 2235, Ap
3	1943	99.0	371	14	US-10-146-502-2235	Sequence 2235, Ap
4	1943	99.0	371	14	US-10-066-543-1422	Sequence 1422, Ap
5	1929	98.3	371	14	US-10-301-822-16	Sequence 16, Appl
6	1929	98.3	371	14	US-10-235-027-126	Sequence 126, App
7	1929	98.3	391	13	US-10-046-935-2239	Sequence 2239, Ap
8	1929	98.3	391	14	US-10-146-502-2239	Sequence 2239, Ap
9	1929	98.3	407	9	US-09-925-301-1165	Sequence 1165, Ap
10	1929	98.3	407	14	US-10-106-698-5654	Sequence 5654, Ap
11	1879.5	95.8	450	14	US-10-301-822-14	Sequence 14, Appl
12	1656	84.4	347	9	US-09-851-588-4	Sequence 4, Appl
13	1656	84.4	347	14	US-10-295-027-1194	Sequence 2, Appl
14	1520.5	77.5	320	9	US-09-851-588-2	Sequence 19, Appl
15	1512.5	77.1	320	14	US-10-264-820-19	Sequence 3438, Ap
16	752	38.3	420	15	US-10-108-260A-3438	Sequence 174871, Ap
17	300.5	15.3	302	15	US-10-424-599-174871	Sequence 272304, Ap
18	299.5	15.3	372	15	US-10-424-599-272304	Sequence 106508, Ap
19	293	14.9	432	16	US-10-437-963-106508	Sequence 274869, Ap
20	258.5	13.2	578	15	US-10-425-114-38958	Sequence 123107, Ap
21	256.5	13.1	220	15	US-10-424-599-274869	Sequence 187737, Ap
22	226	11.5	549	16	US-10-437-963-123107	Sequence 261675, Ap
23	221	11.3	576	16	US-10-437-963-187737	Sequence 46733, A
24	204.5	10.4	221	15	US-10-424-599-261675	Sequence 104286, Ap
25	196	10.0	240	15	US-10-425-114-46733	Sequence 38290, A
26	162.5	8.3	670	16	US-10-437-963-104286	Sequence 210074, A
27	158	8.1	817	15	US-10-425-114-38290	Sequence 38875, A
28	147	7.5	941	15	US-10-424-599-210074	Sequence 38701, A
29	146	7.4	676	15	US-10-425-114-38875	Sequence 878, App
30	138.5	7.1	747	15	US-10-425-114-38701	Sequence 166583, Ap
31	138	7.0	786	14	US-10-225-066A-878	Sequence 254, App
32	138	7.0	786	15	US-10-374-780A-2288	Sequence 204478, Ap
33	131	6.7	225	15	US-10-424-599-166583	Sequence 85, Appl
34	129	6.6	754	14	US-10-153-668-254	Sequence 181788, Ap
35	127	6.5	448	14	US-10-437-963-370	Sequence 229147, Ap
36	122.5	6.2	694	16	US-10-437-963-204478	Sequence 17, Appl
37	115	5.9	422	14	US-10-060-255-85	Sequence 2231, Ap
38	113	5.8	997	16	US-10-437-963-181788	Sequence 147, App
39	111.5	5.7	927	15	US-10-424-599-229147	Sequence 7, Appl
40	111.5	5.7	1795	8	US-08-973-363-17	Sequence 57426, A
41	111.5	5.7	2263	16	US-10-408-765A-2231	Sequence 185, App
42	111	5.7	333	9	US-09-771-161A-147	
43	111	5.7	19652	15	US-10-084-846A-7	
44	110	5.6	354	15	US-10-425-114-57426	
45	110	5.6	536	10	US-09-866-050A-185	

US-09-878-178-2235									
Alignment Scores:									
Pred. No.:	3.06e-178	Length:	371						
Score:	1943.00	Matches:	371						
Percent Similarity:	100.00%	Conservative:	0						
Best Local Similarity:	100.00%	Mismatches:	0						
Query Match:	99.03%	Indels:	0						
DB:	9	Gaps:	0						
US-10-046-935-2234 (1-1116) x US-09-878-178-2235 (1-371)									
QY	1	ATGACGCTCGCGGTCGCGAGAGATCTCAGAGTAAAGAACTTAAAGAAATTC	60						
Db	1	MetAspAlaArgValProGlnLysAspLeuArgValLysLysAsnLeuLysPhe	20						
QY	61	AGATATGTGAAGTTGATTTCCATGGAACCTCTCATCTCTGATGACAGCTTGACAGC	120						
Db	21	ArgTyrValLysLeuLeuSerMetGluThrSerSerSerAspAspSerCysAspSer	40						
QY	121	TTTGCTTCTGATTAATTTGCAACACAGAGCTGCAGTCAGTTCGGGAGGCTTAGGACC	180						
Db	41	PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr	60						
QY	181	CGCAGCAGTGCAGGCACTCTGGACCTCTCAGGTGCGATGAAGTTTCCAGCGCGGAGT	240						
Db	61	ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer	80						
QY	241	ACCAGGGGACCAACCAAAAAAGCAGAGTCCCGCCAGCCCTCAGAGAAATTCGTGACT	300						
Db	81	ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr	100						
QY	301	GATTCGAATCCGATTCAGAGATGAAGTGGATGAATTTTGGAGAAAAGGGCTTTA	360						
Db	101	AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu	120						
QY	361	AATATAAGCAAAACAAAGCAATCTTGCAAAACCTCATGTCTGAATTAGAAAGTCCCT	420						
Db	121	AsnLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro	140						
QY	421	GGCTCGTTCCGTGGAAGACATCCCTCCAGGCTCCGACTCCACATCAAGGAGACCGGA	480						
Db	141	GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg	160						
QY	481	AGCGGTACATTCGGGTGTTGCTTCCAGGAGAAACCTGAAACGAGAGCTCGTCTCTT	540						
Db	161	ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgArgAlaArgProLeu	180						
QY	541	ACCAGGTCAAGTCCGAGATCTCGGGTCCCTTCAGCTCTACCCATGGRGAGAGCCAGT	600						
Db	181	ThrArgSerArgSerArgHisLeuGlySerLeuAspAlaLeuProMet**ArgAlaSer	200						
QY	601	TTKGARKMTMGTCATGTTGGTGAAGAAGAGAGACCGTGGATGGCTACATGAATGAA	660						
Db	201	***Glu***TyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu	220						
QY	661	GATGACCTCCAGAGACCGTCCGATCATCCGTGACCTCCGTCATATAATTCGC	720						
Db	221	AspAspLeuProArgSerArgSerArgSerValThrLeuProHisLilleArg	240						
QY	721	CCAGTGGAGAATAATACAGAGAGAGTGGAGAACGTCCTCAGCAATTCGAGAGAG	780						
Db	241	ProValGluGluLleThrGluGluLeuGluAsnValCysSerAsnSerArgGluLys	260						
QY	781	ATATATAACGTTCACTGGGCTTACTTGTCTCATCAATGCGTCAGAGAACTATTGATCC	840						
Db	261	IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr	280						
QY	841	AAAAAACTCGAGAAACCCAGACTGCTGGGCGGTTTCAGAGCCAGTTCGTGGCCCTTCG	900						
Db	281	LysThrAsnCysArgAsnProAspCysIleTrpGlyValArgGlyGlnPheCysGlyProCys	300						
QY	901	CTTCGAAAACCGTTATGGTGAAGAGGTTCAGGATGCTCTCTGGATCCGAACTGGCATTCG	960						

Db	301	LeuArgAsnArgTyrGlyGluValArgAspAlaLeuLeuAspProAsnTrpHisCys	320
QY	961	CGCGCTTGTGAGGAATCTGCAACTGCACTTCTCGCGCAGCAGATGACCGGTGTGCG	1020
Db	321	ProProCysArgGlyLysCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla	340
QY	1021	ACTGGGGTCTTGTGTATTATTAGCCAAATATCATCGCTTTGGGAATGTGCATGCTACTTG	1080
Db	341	ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu	360
QY	1081	AAAAGCTGAAACAGGAATTTGAAATGCAAGCA	1113
Db	361	LysSerLeuLysGlnGluPheGluMetGlnAla	371
RESULT 2			
US-10-046-935-2235			
; Sequence 2235, Application US/10046935			
; Publication No. US20020156011A1			
; GENERAL INFORMATION:			
; APPLICANT: Jiang, Yugu			
; APPLICANT: Harlocker, Susan L.			
; APPLICANT: Secrist, Heather			
; APPLICANT: Wang, Aijun			
; APPLICANT: Stolk, John A.			
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY			
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER			
; FILE REFERENCE: 210121.527C1			
; CURRENT APPLICATION NUMBER: US/10/046,935			
; CURRENT FILING DATE: 2002-01-15			
; NUMBER OF SEQ ID NOS: 2239			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 2235			
; LENGTH: 371			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: VARIANT			
; LOCATION: 197, 201, 203, 204			
; OTHER INFORMATION: Xaa = Any Amino Acid			
US-10-046-935-2235			
Alignment Scores:			
Pred. No.:	3.06e-178	Length:	371
Score:	1943.00	Matches:	371
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	99.03%	Indels:	0
DB:	13	Gaps:	0
US-10-046-935-2234 (1-1116) x US-10-046-935-2235 (1-371)			
QY	1	ATGACGCTCGCGGTCGCGAGAGATCTCAGAGTAAAGAACTTAAAGAAATTC	60
Db	1	MetAspAlaArgValProGlnLysAspLeuArgValLysLysAsnLeuLysPhe	20
QY	61	AGATATGTGAAGTTGATTTCCATGGAACCTCTCATCTCTGATGACAGCTTGACAGC	120
Db	21	ArgTyrValLysLeuLeuSerMetGluThrSerSerSerAspAspSerCysAspSer	40
QY	121	TTTGCTTCTGATTAATTTGCAACACAGAGCTGCAGTCAGTTCGGGAGGCTTAGGACC	180
Db	41	PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr	60
QY	181	CGCAGCAGTGCAGGCACTCTGGACCTCTCAGGTGCGATGAAGTTTCCAGCGCGGAGT	240
Db	61	ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer	80
QY	241	ACCAGGGGACCAACCAAAAAAGCAGAGTCCCGCCAGCCCTCAGAGAAATTCGTGACT	300
Db	81	ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr	100
QY	301	GATTCGAATCCGATTCAGAGATGAAGTGGATGAATTTTGGAGAAAAGGGCTTTA	360
Db	101	AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu	120
QY	361	AATATAAGCAAAACAAAGCAATCTTGCAAAACCTCATGTCTGAATTAGAAAGTCCCT	420
Db	121	AsnLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro	140
QY	421	GGCTCGTTCCGTGGAAGACATCCCTCCAGGCTCCGACTCCACATCAAGGAGACCGGA	480
Db	141	GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg	160
QY	481	AGCGGTACATTCGGGTGTTGCTTCCAGGAGAAACCTGAAACGAGAGCTCGTCTCTT	540
Db	161	ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgArgAlaArgProLeu	180
QY	541	ACCAGGTCAAGTCCGAGATCTCGGGTCCCTTCAGCTCTACCCATGGRGAGAGCCAGT	600
Db	181	ThrArgSerArgSerArgHisLeuGlySerLeuAspAlaLeuProMet**ArgAlaSer	200
QY	601	TTKGARKMTMGTCATGTTGGTGAAGAAGAGAGACCGTGGATGGCTACATGAATGAA	660
Db	201	***Glu***TyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu	220
QY	661	GATGACCTCCAGAGACCGTCCGATCATCCGTGACCTCCGTCATATAATTCGC	720
Db	221	AspAspLeuProArgSerArgSerArgSerValThrLeuProHisLilleArg	240
QY	721	CCAGTGGAGAATAATACAGAGAGAGTGGAGAACGTCCTCAGCAATTCGAGAGAG	780
Db	241	ProValGluGluLleThrGluGluLeuGluAsnValCysSerAsnSerArgGluLys	260
QY	781	ATATATAACGTTCACTGGGCTTACTTGTCTCATCAATGCGTCAGAGAACTATTGATCC	840
Db	261	IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr	280
QY	841	AAAAAACTCGAGAAACCCAGACTGCTGGGCGGTTTCAGAGCCAGTTCGTGGCCCTTCG	900
Db	281	LysThrAsnCysArgAsnProAspCysIleTrpGlyValArgGlyGlnPheCysGlyProCys	300
QY	301	GATTCGAATCCGATTCAGAGATGAAGTGGATGAATTTTGGAGAAAAGGGCTTTA	360

```
Db 101 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 120
QY 361 AATATAAAGCAAAACAAAGCAATGCTTGCAAACTCATGCTCTGAATTAAGAAAGCTCCCT 420
Db 121 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140
QY 421 GGCTCGTTCCGTGGGAAGACATCCCTCCAGGCTCCGACTCAATCAAGGAGACCGGA 480
Db 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 160
QY 481 AGCGTACATCCCGGTGTTGCTTCCAGGAGAAACCTGACGAGAGCTGCTCTT 540
Db 161 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgAlaArgProLeu 180
QY 541 ACCAGGTCAAGTCCCGGATCTCCGGTCCCTTGACGCTCTACCCATGGRGAGCCAGT 600
Db 181 ThrArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMet**ArgAlaSer 200
QY 601 TTGARKMTWGTACATGTTGGTGAAGAAAGAGAGACCGTGGATGGCTACATGAATGAA 660
Db 201 ***Glu****TyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 220
QY 661 GATGACCTGCCAGAGCCGCTCCAGATCATCCGTCAGCCCTTCCGCATATAATTCG 720
Db 221 AspAspLeuProArgSerArgSerArgSerValThrLeuProHisIleIleArg 240
QY 721 CCAGTGAAGAAATACAGAGGAGAGTTGGAGAACGCTCTCAGCAATTCCTCGAGAGAG 780
Db 241 ProValGluGluIleThrGluGluGluLeuGluAsnValCysSerAsnSerArgGluLys 260
QY 781 ATATATAACCGTTCACTGGGCTCTACTTGTCAATCAATGCCGTGAGAGACTATTGATAAC 840
Db 261 IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 280
QY 841 AAAAACAACAGCAAAACCCAGACTGCTGGGCGCTTGAGGCCAGTCTGTGGCCCTTCG 900
Db 281 LysThrAsnCysArgAsnProAspCysTrpGlyValArgGlyGlnPheCysGlyProCys 300
QY 901 CTTGCAACCGTTATGCTGAAGAGCTCAGGAGATCGTCTGCTGGATCCGACTGCGCATTCG 960
Db 301 LeuArgAsnArgTyrGlyGluGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 320
QY 961 CCGCTTGTGAGAAATCTCAACTGCACTGCTTCTGCCGAGCAGAGATGACCGTGTGCG 1020
Db 321 ProProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 340
QY 1021 ACTGGGCTCTGTGATTTAGCCMAATATCATGCTTGGGAATGTCATGCTACTTG 1080
Db 341 ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 360
QY 1081 AAAAGCTGAACAGGAATTTGAAATGCAAGCA 1113
Db 361 LysSerLeuLysGlnGluPheGluMetGlnAla 371
```

RESULT 3

```
US-10-146-502-2235
; Sequence 2235, Application US/10146502
; Publication No. US20030069180A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aijun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.527C2
; CURRENT APPLICATION NUMBER: US/10/146,502
; CURRENT FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2241
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 2235
```

```
; LENGTH: 371
; TYPE: PRF
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 197, 201, 203, 204
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-146-502-2235

Alignment Scores:
Pred. No.: 3,06e-178 Length: 371
Score: 1943.00 Matches: 371
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.03% Indels: 0
DB: 14 Gaps: 0

US-10-046-935-2234 (1-1116) x US-10-146-502-2235 (1-371)
```

```
QY 1 ATGGACGCTCCCGGTCGCCGCAAGAAAGATCTCAGAGTAAAGAAAGAACTTAAAGAAATTC 60
Db 1 MetAspAlaArgArgValProGlnLysAspLeuArgValLysLysAsnLeuLysLysPhe 20
QY 61 AGATATGTGAAGTTGATTTCATCGAAACCTCGTCATCCTCTGATGACAGTTGTGACAGC 120
Db 21 ArgTyrValLysLeuIleSerMetGluThrSerSerSerSerSerSerSerSerSerSer 40
QY 121 TTTCTCTTGTATAATTTTGCACACACAGAGGTCGAGTCAGTTCGGGAAGGCTGTAGGACC 180
Db 41 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr 60
QY 181 CGCAGCCAGTGCAGGCACTCTGGACCTCTCAGGGTGGCGATGAAGTTTCCAGCCGGAGT 240
Db 61 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 80
QY 241 ACCAGGGGAGCAACCAACAAAGAGAGAGTCCCGCCAGCCCTCAGAGAAATCTGTGACT 300
Db 81 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluLysSerValThr 100
QY 301 GATTCCAACTCCGATTCAGAAAGATGAAGTGAATGAATTTTGGAGAAAAGGCTTTA 360
Db 101 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 120
QY 361 AATATAAAGCAAAACAAAGCAATGCTTGCAAACTCATGCTCTGAATTAAGAAAGCTCCCT 420
Db 121 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140
QY 421 GGCTCGTTCCGTGGAAGACATCCCTCCAGGCTCCGACTCACAATCAAGGAGACCGGA 480
Db 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 160
QY 481 AGCGTACATCCCGGTGTTGCTTCCAGGAGAAACCTGACGAGAGCTCGCTCTT 540
Db 161 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgAlaArgProLeu 180
QY 541 ACCAGGTCAAGTCCCGGATCTCCGGTCCCTTGACGCTCTACCCATGGRGAGCCAGT 600
Db 181 ThrArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMet**ArgAlaSer 200
QY 601 TTGARKMTWGTACATGTTGGTGAAGAAAGAGACCGTGGATGGCTACATGAATGAA 660
Db 201 ***Glu****TyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 220
QY 661 GATGACCTGCCAGAGCCGCTCCAGATCATCCGTCAGCCCTTCCGCATATAATTCG 720
Db 221 AspAspLeuProArgSerArgSerArgSerValThrLeuProHisIleIleArg 240
QY 721 CCAGTGAAGAAATTAAGAGAGAGTTGGAGAACGCTCTGAGCAATTCCTCGAGAGAG 780
Db 241 ProValGluGluIleThrGluGluGluLeuGluAsnValCysSerAsnSerArgGluLys 260
QY 781 ATATATAACCGTTCACTGGGCTCTACTTGTCAATCAATGCCGTGAGAGACTATTGATAAC 840
```

Db 261 lletyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 280
QY 841 AAAACAACTGCAGAAACCCAGACTGCTGGGGCGTTCGAGGCCAGTTCGTGCGCCCTGC 900
Db 281 LysThrAsnCysArgAsnProAspCysTrpGlyValArgGlyGlnPheCysGlyProCys 300
QY 901 CTTCCGAACCGTTATGGTGAAGAGGTCAGGGATGCTCTCTGGATCCGAACTGGCATTGC 960
Db 301 LeuArgAsnArgTrpGlyGlnGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 320
QY 961 CGCCCTTGTCCGAGGAATTCGCAACTGCACTTCTGCCGCGCAGCAGATGAGCGGTGCG 1020
Db 321 ProProCysArgGlyIleCysAsnCysSerPheCysArgGluArgAspGlyArgCysAla 340
QY 1021 ACTGGGTCCTGTGTATTAGCCAAATATCATGGCTTTGGGNAATGTGATGCTACTTG 1080
Db 341 ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 360
QY 1081 AAAAGCCTGAAACAGGAATTTGAAATGCAAGCA 1113
Db 361 LysSerLeuLysGlnLysPheGluMetGlnAla 371

RESULT 4

US-10-066-543-1422

; Sequence 1422, Application US/10066543

; Publication No. US20030087818A1

; GENERAL INFORMATION:

; APPLICANT: Jiang, Yuqiu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margareta
; APPLICANT: Stolk, John A.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.563

; CURRENT APPLICATION NUMBER: US/10/066,543

; CURRENT FILING DATE: 2002-01-31

; NUMBER OF SEQ ID NOS: 3417

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1422

; LENGTH: 371

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: VARIANT

; LOCATION: 197,201,203,204

; OTHER INFORMATION: Xaa = Any amino acid

US-10-066-543-1422

Alignment Scores:
Pred. No.: 3,06e-178 Length: 371
Score: 1943.00 Matches: 371
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.03% Indels: 0
DB: 14 Gaps: 0

US-10-046-935-2234 (1-1116) x US-10-066-543-1422 (1-371)

QY 1 ATGACGCTCGCCGTCGCGCAGAAAGATCTCAGAGTAAGAGAACTTAAGAAATTC 60
Db 1 MetAspAlaArgValProGlnLysAspLeuArgValLysLysAsnLeuLysPhe 20
QY 61 AGATATGTGAAGTTCATTTCCATGGAACCTCGTCATCTCTGATGACAGTGTGACAGC 120
Db 21 ArgTyrValLysLeuLeuSerMetGluThrSerSerSerAspSerCysAspSer 40

QY 121 TTTGCTTCTGATAATTTTGCACAAACAGAGGCTGCAGTCAGTTCCGGAAAGCTGTAGGACC 180
Db 41 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGlnGlyCysArgThr 60
QY 181 CGACGCCAGTGCAGACACTCTCGACTCTCAGGGTGGCGATGAAGTTTTCAGCGCGGAGT 240
Db 61 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 80
QY 241 ACCAGGGAGCAACCAACAAAAAGCAGAGTCCCGCCAGCCCTCAGAGAAATTCGTGACT 300
Db 81 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 100
QY 301 GATTCAACTCCGATTCAGAACTGAAAGTGAATCAATTTTTCGAGAAAGGGCTTTTA 360
Db 101 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGlnLysArgAlaLeu 120
QY 361 ATATATAAGCAAAACAAACCAATCTTGCAAAACATCATGCTGAAATAGAAAAGCTTCCCT 420
Db 121 AsnLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140
QY 421 GGCTCGTTCGGTGAAGACATCCCTCCAGGCTCCGACTCACAATCAAGGACACCGCGA 480
Db 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 160
QY 481 AGCGTACATTCGCGGTGTGCTTCAGGAGAAACCCCTGAACGAGAGCTCGTCCCTCT 540
Db 161 ArgArgThrPheProGlyValAlaSerArgAsnProGluArgArgAlaArgProLeu 180
QY 541 ACCAGGTCAAGTCCCGGATCCCTCGGCTCCCTGAGGCTCTACCCATGGRGAGAGCCAGT 600
Db 181 ThrArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMet***ArgAlaSer 200
QY 601 TTXGARMTMWGTACATGTTGGTGAGAAAGAGAACCCGTGGATGGCTACATGAATGAA 660
Db 201 ***Glu***TyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 220
QY 661 GATGACCTGCCAGAGCCGCTCGCTCAGATCATCCGTGACCTTCGCGATATAATTCGC 720
Db 221 AspAspLeuProArgSerArgArgSerValThrLeuProHisIleLeuArg 240
QY 721 CCAGTGAAGAAATACAGAGAGGAGTTCGAGAAACGCTCTGCAGCAATTCGAGAGAAG 780
Db 241 ProValGluGlnIleThrGluGluGluLeuAsnValCysSerAsnSerArgGluLys 260
QY 781 ATATATAACCGTTCACTGGCTCTACTTGTCAATCAATGCCGTGAGAGACTATTGATACC 840
Db 261 lletyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 280
QY 841 AAAACAACTGCAGAAACCCAGACTGCTGGGGCGTTCGAGCCAGTTCTGTGGCCCTGC 900
Db 281 LysThrAsnCysArgAsnProAspCysTrpGlyValArgGlyGlnPheCysGlyProCys 300
QY 901 CTTCCGAACCGTTATGGTGAAGAGTTCAGGGATGCTCTGCTGGATCCGAACTGGCATTCG 960
Db 301 LeuArgAsnArgTyrGlyGluGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 320
QY 961 CGCCTTGTTCAGGAATCTGCAACTGCAGTTTCTGCCGCGCAGCAGATGAGCGGTGTCGG 1020
Db 321 ProProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 340
QY 1021 ACTGGGTCCTGTGTATTAGCCAAATATCATGCTTTGGAAATGTGCATGCCCTACTTG 1080
Db 341 ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 360
QY 1081 AAAAGCCTGAAACAGGAATTTGAAATGCAAGCA 1113
Db 361 LysSerLeuLysGlnLysPheGluMetGlnAla 371

RESULT 5

US-10-301-822-16

; Sequence 16, Application US/10301822

; Publication No. US20030148410A1

; GENERAL INFORMATION:

APPLICANT: Millennium Pharmaceuticals, Inc.
APPLICANT: Berger, Allison
APPLICANT: Guillemette, Tracy L.
APPLICANT: Kamatkar, Shubhangi
APPLICANT: Schlegel, Robert
APPLICANT: Monahan, John E.
APPLICANT: Thibodeau, Stephen N.
APPLICANT: Burgart, Lawrence J.
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
THERAPY OF COLON CANCER
FILE REFERENCE: MP001-029P2RNM
CURRENT APPLICATION NUMBER: US/10/301,822
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 60/339,971
PRIOR FILING DATE: 2001-12-10
PRIOR APPLICATION NUMBER: US 60/361,978
PRIOR FILING DATE: 2002-03-05
PRIOR APPLICATION NUMBER: US 60/381,989
PRIOR FILING DATE: 2002-05-20
NUMBER OF SEQ ID NOS: 228
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16
LENGTH: 371
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-301-822-16

Alignment Scores:
Pred. No.: 6.87e-177 Length: 371
Score: 1929.00 Matches: 364
Percent Similarity: 98.11% Conservative: 0
Best Local Similarity: 98.11% Mismatches: 7
Query Match: 98.32% Indels: 0
DB: 14 Gaps: 0

US-10-046-935-2234 (1-1116) x US-10-301-822-16 (1-371)

QY 1 ATGACGCTCCGCGTCCGCGAGAAAGATCTCAGAGTAAAGAAAGAACTTAAAGAAATTC 60
Db 1 MetAspAlaArgValProGlnLysAspLeuArgValLysLysAsnLeuLysLysPhe 20

QY 61 AGATATGTGAAGTTGATTTCATGGAAACCTCGTCATCCTCTGTATGACAGTTGTGACAGC 120
Db 21 ArgTyrValLysLeuIleSerMetGluThrSerSerSerSerSerSerSerSerSerSer 40

QY 121 TTTGCTTCTCATATTTTGCACACACAGGCTGCAGTCAGTCCGGAAGGCTGTAGGACC 180
Db 41 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr 60

QY 181 CGGAGCCAGTGCAGGCACTCTGGACCTCTCAGGTTGGCGATGAAGTTTCCAGCGCGGAGT 240
Db 61 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 80

QY 241 ACCAGGGGACCAACCAAAAGCAGAGTCCGCGCAGCCCTCAGAGAATTTCTGTGACT 300
Db 81 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 100

QY 301 GATTCCAACCTCCATTGAGAGATGAAGTGAATGAATTTTGGAGAAAGGCTTTA 360
Db 101 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 120

QY 361 AATATAAGCAAAACAAAGCAATCTTTCGCAAACTCATGTCTGAATTAGAAAGCTTCCCT 420
Db 121 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140

QY 421 GGCTCGTTCGTTGGAAGACATCCCTCCAGGCTCCGACTCAATCAAGAGACCGCGA 480
Db 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgProArg 160

QY 481 AGCGGTACATTCGCGGTGTGTCTCCAGGAGAAACCTTGAACGAGAGCTCTCTCTT 540
Db 161 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgArgAlaArgProLeu 180

QY 541 ACCAGGTCAAGTCCCGATCTCTGGGTCCCTTGAGCTCTACCCATGGRGAGAGCCAGT 600
Db 181 ThrArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMetGluGluGlu 200

QY 601 TTKGARKMTWGTACATGTTGGTGAAGAGGAGACCGTGGATGGCTACATGAATGAA 660
Db 201 GluGluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 220

QY 661 GATGACCTGCCCAAGACCGTCCAGATCATCCGTGACCCCTTCCGCATATAATTCGC 720
Db 221 AspAspLeuProArgSerArgArgSerValThrLeuProHisIleIleArg 240

QY 721 CCAGTGGAGAAATACAGAGAGAGATTGGAGAACGTCTGCAGCAATTCGAGAGAG 780
Db 241 ProValGluGluIleThrGluGluGluLeuValCysSerAsnSerArgGluLys 260

QY 781 ATATATAACCGTTCACCTGGGCTCTACTTGTCTCATCAATCCCTCAGAGACTATTGATACC 840
Db 261 IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 280

QY 841 AAAACAAACTGCAGAAACCCAGACTGCTGGGGCGTTCGAGCCAGTTCTGTGGCCCTGC 900
Db 281 LysThrAsnCysArgAsnProAspCysTyrGlyValArgGlyGlnPheCysGlyProCys 300

QY 901 CTTCCGAAACCGTTATGCTGAGAGGTCAGGATCTCTCTGGATCCGAACTGGCATTCG 960
Db 301 LeuArgAsnArgTyrGlyGluGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 320

QY 961 CCGCTTTGTCGAGGAATCTCACTGCAGTTTCTGCCGAGCAGAGATGACGGTGTGCG 1020
Db 321 ProProCysArgGlyLysCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 340

QY 1021 ACTGGGTCTCTGTGTTATTTAGCCAAATATCATGGCTTTGGGATGTCATGCTCTACTTG 1080
Db 341 ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 360

QY 1081 AAAAGCTGAAACAGGAATTTGAAATGCAAGCA 1113
Db 361 LysSerLeuLysGlnGluPheGluMetGlnAla 371

RESULT 6
US-10-295-027-126
Sequence 126, Application US/10295027
Publication No. US20030232350A1
GENERAL INFORMATION:
APPLICANT: Afar, Daniel
APPLICANT: Aziz, Natasha
APPLICANT: Ginsberg, Wendy M.
APPLICANT: Gish, Kurt C.
APPLICANT: Glynn, Richard
APPLICANT: Hevezi, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Murray, Richard
APPLICANT: Watson, Susan R.
APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
Methods of Screening for Modulators of Cancer
FILE REFERENCE: 018501-012500US
CURRENT APPLICATION NUMBER: US/10/295,027
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/663,733
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/335,394
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/332,464
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: US 60/334,393
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: US 60/340,376
PRIOR FILING DATE: 2001-12-14

```
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 126
; LENGTH: 371
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-295-027-126

Alignment Scores:
Pred. No.: 6.87e-177 Length: 371
Score: 1929.00 Matches: 364
Percent Similarity: 98.11% Conservative: 0
Best Local Similarity: 98.11% Mismatches: 7
Query Match: 98.32% Indels: 0
DB: 14 Gaps: 0

US-10-046-935-2234 (1-1116) x US-10-295-027-126 (1-371)

QY 1 ATGGAGCTGCGCGCTGCGGAGAAAGATCTCAGAGTAAGAAGAACTTAAAGAATTC 60
Db 1 MetAspAlaArgValProGlnLysAspLeuArgValLysLysAsnLeuLysLysPhe 20

QY 61 AGATATGTGAAGTTGATTTCCATGGAACCTCGTCATCTCATGACAGTGTGACAGC 120
Db 21 ArgTyrValLysLeuLysSerMetGluThrSerSerSerSerSerSerSerSerSer 40

QY 121 TTTGCTTCTGATTAATTTTGCACACAGAGCTCGAGTCAGTCAGTTCGGGAAGCTGTAGAAC 180
Db 41 PheAlaSerAspAsnPheAlaAenThrArgLysGlnSerValArgGluGlyCysArgThr 60

QY 181 CGAGCGAGTGCAGGCACTCGACCTTCAGGTCGGATGAAGTTTCAGCGCGGAGT 240
Db 61 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 80

QY 241 ACCAGGGAGCAACCAACAAAGACAGAGTCCCGCCAGCCCTCAGAGAAATTCGTGACT 300
Db 81 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 100

QY 301 GATTCGAATCCGATTCAGAGATGAAGTGGAAATGAATTTTGGAGAAAAGGGCTTTA 360
Db 101 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 120

QY 361 AATATAAGCAAAACAAAGCAATCTTGCACAACTCATGCTGAATTAGAAGCTTCCT 420
Db 121 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140

QY 421 GGCTCGTTTCGTGGAAGACATCCCTCCAGGCTCCGACTCACAAATCAAGAGACCGGA 480
Db 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 160

QY 481 AGCGGTACATTCGCGGTTGTTGTTCCAGGAGAAACCCCTGAACCGAGAGCTCGTCTTT 540
Db 161 ArgArgThrPheProGlyValAlaSerArgAsnProGluArgArgAlaArgProLeu 180

QY 541 ACCAGGTCAGGTCAGGATCCTCGGTCCTTCAGCGCTCTACCATCGRGAGAGCCAGT 600
Db 181 ThrArgSerArgSerArgGlyLeuGlySerLeuAspAlaLeuProMetGluGluGlu 200

QY 601 TTKGARKMTWMTACATGTTGTGAGAAAGAGAACCGGTGGATGGCTACATGAATGAA 660
Db 201 GluGluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 220

QY 661 GATGACCTGCCAGAACCGGTGCTCCAGATCATCCGTGACCCCTCCGCATATAATTCG 720
Db 121 TTTGCTTCTGATTAATTTTGCACACAGAGCTCGAGTCAGTTCGGGAAGCTGTAGAAC 180

221 AspAspLeuProArgSerArgArgSerArgSerValThrLeuProHisIleLeuArg 240
721 CCAGTGGAGAAATACAGAGGAGGAGTTGGAGAACGCTCGACCAATTCGAGAGAAG 780
241 ProValGluGluIleThrGluGluLeuGluAsnValCysSerAsnSerArgGluLys 260
781 ATATATAACCGTTCACTGGGCTCTACTTGTCTCATCAATGCCGTCAGAGACTATTGATACC 840
261 IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 280
841 AAAACAAACTGCAGAAACCCAGACAGCTGCTGGGGGCTTCGAGGCCAGTTCGTGGCCCTGC 900
281 LysThrAsnCysArgAsnProAspCysTrpGlyValArgGlyGlnPheCysGlyProCys 300
901 CTTGAAACCGTTATGCTGAAGAGTCAAGGATGCTCTCTGTCGATCCGAACCTGCATTGC 960
301 LeuArgAsnArgTyrGlyGluValAlaArgAspAlaLeuLeuAspProAsnTrpHisCys 320
961 CCGCCTTGTGAGGAATCTGCAACTCAGTTCCTGCGGCGAGAGATGCGGCTGTGCG 1020
321 ProProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 340
1021 ACTGGGCTCCTGTGTATTTAGCCAAATATCATGGCTTTGGGAATGTGCATGCTACTTTC 1080
341 ThrGlyValLeuValLysLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 360

1081 AAAAGCCTGAACAGGAATTTGAAATGCAAGCA 1113
361 LysSerLeuLysGlnGluPheGluMetGlnAla 371

RESULT 7
US-10-046-935-2239
; Sequence 2239, Application US/10046935
; Publication No. US20020156011A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aijun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.527C1
; CURRENT APPLICATION NUMBER: US/10/046,935
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 2239
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 2239
; LENGTH: 391
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-046-935-2239

Alignment Scores:
Pred. No.: 7.02e-177 Length: 391
Score: 1929.00 Matches: 364
Percent Similarity: 98.11% Conservative: 0
Best Local Similarity: 98.11% Mismatches: 7
Query Match: 98.32% Indels: 0
DB: 13 Gaps: 0

US-10-046-935-2234 (1-1116) x US-10-046-935-2239 (1-391)

QY 1 ATGGAGCTGCGCGCTGCGGAGAAAGATCTCAGAGTAAGAAGAACTTAAAGAATTC 60
Db 21 MetAspAlaArgValProGlnLysAspLeuArgValLysLysAsnLeuLysLysPhe 40

QY 61 AGATATGTGAAGTTGATTTCCATGGAACCTCGTCATCTCATGACAGTGTGACAGC 120
Db 41 ArgTyrValLysLeuLysSerMetGluThrSerSerSerSerSerSerSerSerSer 60

121 TTTGCTTCTGATTAATTTTGCACACAGAGCTCGAGTCAGTTCGGGAAGCTGTAGAAC 180
```



```
Db 61 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr 80
QY 181 CGCAGCCAGTCAGGCACTCTGGACCTCTCAGGCTGGCGATGAAGTATTTCCAGCGCGGAGT 240
Db 81 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 100
QY 241 ACCAGGGAGCAACCAACAAAAAGCAGAGTCGCCCGCCAGCCCTCAGAGAAATTTCTGTGACT 300
Db 101 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 120
QY 301 GATTCCCACTCCGATTCAGAGATCAAAAGTGAATGAATTTTGGAGAAAAAGCGCTTTA 360
Db 121 AspSerAsnSerAspSerGlyAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 140
QY 361 AATATAAGCAAAAACAAAGCAATGCTTCAAAACCTCATGTCTGAATTAAGAAAGTTCCTCCT 420
Db 141 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 160
QY 421 GGCTCGTTCGGTGAAGACATCCCTCCAGGCTCCGACTCACAAATCAAGGAGACCGGGA 480
Db 161 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 180
QY 481 AGCGGTACATTCGCGGTGTTGCTTCAGAGGAAACCTGAAACCGAGAGCTCGTCTCTT 540
Db 181 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgAlaArgProLeu 200
QY 541 ACCAGGTCAAAGTCCCGGATCCTCGGGTCCCTTCAGCGCTCTACCCATGGRGAGAGCCAGT 600
Db 201 ThrArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMetGluGluGlu 220
QY 601 TTKGARMTMWGTACATGTTGGTGAGAAAGAGACCGGTGGATGGCTACATGAATGAA 660
Db 221 GluGluAspLysIleMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 240
QY 661 GATGACCTGCCAGAACCGTCGCTCCAGATCATCGGTGACCCCTTCGCGATATTAATTCG 720
Db 241 AspAspLeuProArgSerArgSerArgSerArgSerValThrLeuProHisIleArg 260
QY 721 CCAGTGAAGAATAATACAGAGGAGGAGTGGAGAACGCTCTGCAGCAATTTCTCGAGAGAG 780
Db 261 ProValGluGluIleThrGluGluGluLeuGluAsnValCysSerAsnSerArgGluLys 280
QY 781 ATATATAACCGTTCACCTGGGCTTACTTGTGCATCAATGCCGTGAGAGACTATTGATACC 840
Db 281 IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 300
QY 841 AAAACAACCTGCAGAACCCAGACTGCTGGGCGCTTCAGGCCAGTTCGTGGCCCTCGC 900
Db 301 LysThrAsnCysArgAsnProAspCysIlePglValArgGlyGlnPheCysGlyProCys 320
QY 901 CTTTCAAAACCGTTATGTTGAAGAGGTCCAGGATGCTCTGTGGATCCGAACCTGSCATTGC 960
Db 321 LeuArgAsnArgTyrGlyGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 340
QY 961 CGCGCTTGTGAGAGATCTGCAACTGCACTGTTTGGCGGAGGAGATGACCGTGTGCG 1020
Db 341 ProProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 360
QY 1021 ACTGGGCTCCTGTGATTTTAGCCAAATATCATGGCTTTGGGATGTGCACTGCTTCTG 1080
Db 361 ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 380
QY 1081 AAAAGCCTGAAACAGGAATTTGAAATCAAGCA 1113
Db 381 LysSerLeuLysGlnGluPheGluMetGlnAla 391
```

RESULT 8

```
US-10-146-502-2239
; Sequence 2239, Application US/10146502
; Publication No. US20030069180A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuciu
; APPLICANT: Harlocker, Susan L.
```

```
; APPLICANT: Secretist, Heather
; APPLICANT: Wang, Aijun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; CURRENT FILING DATE: 2002-05-14
; CURRENT APPLICATION NUMBER: US/10/146,502
; NUMBER OF SEQ ID NOS: 2241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2239
; LENGTH: 391
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-146-502-2239
Alignment Scores:
Pred. No.: 7,02e-177 Length: 391
Score: 1929,00 Matches: 364
Percent Similarity: 98.11% Conservative: 0
Best Local Similarity: 98.11% Mismatches: 7
Query Match: 98.32% Indels: 0
DB: 14 Gaps: 0
US-10-046-935-2234 (1-1116) x US-10-146-502-2239 (1-391)
```

```
QY 1 ATGACACCTCGCCGCTGCGCAGAAAGATCTCAGAGTAAGAAGAACTTAAAGAAATTC 60
Db 21 MetAspAlaArgArgValProGlnLysAspLeuArgValLysAsnLeuLysLysPhe 40
QY 61 AGATATCTGAAGTTGATTTCCATGGAACCTCGTCATCTCTGATGACAGTTGTGACAGC 120
Db 41 ArgTyrValLysLeuIleSerMetGluThrSerSerSerSerSerSerSerSerSer 60
QY 121 TTTGCTTCTGATATTTTGCAACACACGAGCTGCAGTCAGTTCGGGAGGCTGTAGGACC 180
Db 61 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr 80
QY 181 CGCAGCCAGTCAGGCACTCTGGACCTCTCAGGCTGGCGATGAAGTATTTCCAGCGCGGAGT 240
Db 81 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 100
QY 241 ACCAGGGAGCAACCAACAAAAAGCAGAGTCGCCCGCCAGCCCTCAGAGAAATTTCTGTGACT 300
Db 101 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 120
QY 301 GATTCCCACTCCGATTCAGAGATGAAGTGAATGAATTTTGGAGAAAAAGCGCTTTA 360
Db 121 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 140
QY 361 AATATAAGCAAAAACAAAGCAATGCTTGGCAAACTCATGTCTGAATTAGAAAGCTTCCTC 420
Db 141 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 160
QY 421 GGCTCGTTCGGTGAAGACATCCCTCCAGGCTCCGACTCACAAATCAAGGAGACCGGGA 480
Db 161 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 180
QY 481 AGCGGTACATTCGCGGTGTTGCTTCAGAGGAAACCTGAAACCGAGAGCTCGTCTCTT 540
Db 181 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgAlaArgProLeu 200
QY 541 ACCAGGTCAAAGTCCCGGATCCTCGGGTCCCTTCAGCGCTCTACCCATGGRGAGAGCCAGT 600
Db 201 ThrArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMetGluGluGlu 220
QY 601 TTKGARMTMWGTACATGTTGGTGAGAAAGAGACCGGTGGATGGCTACATGAATGAA 660
Db 221 GluGluAspLysIleMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 240
QY 661 GATGACCTGCCAGAACCGTCGCTCCAGATCATCGGTGACCCCTTCGCGATATTAATTCG 720
Db 241 AspAspLeuProArgSerArgSerArgSerArgSerValThrLeuProHisIleArg 260
```



```
QY 721 CCAGTGGAGAAATTACAGCAGGAGGTTGGAGAACGTTCTGCAGCAATTTCTCAGAGAAG 780
Db 261 ProValGluGluLeuThrGluGluLeuGluValCysSerAsnSerArgGluLys 280
QY 781 ATATATAACCGTTCACCTGGCTCTACTTGTCTCATCAATGCCGTCAGAGACTATTGATACC 840
Db 281 IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 300
QY 841 AAAAACHAATCTCAGAACCCAGACTCTGGGGGTTTCAGAGCCAGTCTGTGGCCCTGC 900
Db 301 LysThrAsnCysArgAsnProAspCysTrpGlyValArgGlyGlnPheCysGlyProCys 320
QY 901 CTTCCGAAACCGTTATCGTGAAGAGTCAGGATGCTCTGTGATCCGAACTGGCATTGC 960
Db 321 LeuArgAsnArgTyrGlyGluGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 340
QY 961 CGCCTTGTTCGAGGAATCTGCAACTGCACTGCTTCTGCCGAGCAGATGACCGTGTGCG 1020
Db 341 ProProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 360
QY 1021 ACTGGGCTCTGTGTATTGTCGAAATATCATGCTTGGGAATGTGCATGCTACTTG 1080
Db 361 ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 380
QY 1081 AAAAGCCTGAAACAGCAATTTGAAATGCAAGCA 1113
Db 381 LysSerLeuLysGlnGluPheGluMetGlnAla 391

RESULT 9
US-09-925-301-1165
; Sequence 1165, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1165
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1165

Alignment Scores:
Pred. No.: 7,13e-177 Length: 407
Score: 1929.00 Matches: 364
Percent Similarity: 98.11% Conservative: 0
Best Local Similarity: 98.11% Mismatches: 7
Query Match: 98.32% Indels: 0
DB: 9 Gaps: 0

US-10-046-935-2234 (1-1116) x US-09-925-301-1165 (1-407)
QY 1 ATGACGCTCGCCGCGTCCGCGAGAAAGATCTCAGAGTAAGAAAGAACTTAAGAAATTC 60
Db 37 MetAspAlaArgArgValProGlnLysAspLeuArgValLysAsnLeuLysLysPhe 56
QY 61 AGATATGTGAAGTTGATTTCATCGAAACCTCGTCATCCTCTGATGACAGTTGTGACGC 120
Db 57 ArgTyrValLysLeuIleSerMetGluThrSerSerSerSerSerSerSerSerSer 76
QY 121 TTTGCTTCTGTAATTTTGCAAAACACGAGGCTGCAAGTTCAGTTCGGGAGGCTGAGGACC 180
Db 77 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr 96
```

```
QY 181 CCGACGCACTGCGGCACTCTCGACCTCTCAGGGTGGCCGATGAGTTCAGCGCGGACT 240
Db 97 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 116
QY 241 ACCGGGAGACCAACAAACAAAAGCAGAGTCCCGCCAGCCCTCAGAGAAATTCGTGCACT 300
Db 117 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 136
QY 301 GATTCCAACTCCGATTTCAGAGATGAAAGTGAATGAATTTTGGAGAAAAGGGCTTTA 360
Db 137 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 156
QY 361 ATATATAACCAACAAACCAATCTTGCAAAACCTCATGCTCAATAGAAAGCTTCCCT 420
Db 157 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 176
QY 421 GGTCTGTCCTCGTGAAGACATCCCTCCAGGCTCCGACTCACAATCAAGAGACACGCGA 480
Db 177 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 196
QY 481 AGCGCTACATTCCTCGGCTGTTGCTTCCAGGAGAAACCTTGAAACGGAGAGCTCTCTCTT 540
Db 197 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgAlaArgProLeu 216
QY 541 ACCAGTCAAGTCCCGGATCTCTCGGTCCTTGACGCTCTACCCATGGRGAGAGCCAGT 600
Db 217 ThrArgSerArgSerArgIleLeuGlySerLeuAspAlaLeuProMetGluGluGlu 236
QY 601 TTKGARKMTWMGTACATGTTGGTGAGAAAGAGAAAGACCTGCGATGCTACATGAATGAA 660
Db 237 GluGluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 256
QY 661 GATGACCTGCCAGAAAGCGTCTCCAGATCATCCGTGACCCCTTCCGCAATATAATTCGC 720
Db 257 AspAspLeuProArgSerArgArgSerArgSerValThrLeuProHisIleLeuArg 276
QY 721 CCAGTGAAGAAATTAACAGAGGAGGTGGAGAACTCTGCAGCAATTCCTCGAGAGAAG 780
Db 277 ProValGluGluIleThrGluGluGluLeuGluAsnValCysSerAsnSerArgGluLys 296
QY 781 ATATATAACCGTTCACCTGGGCTCTACTTGTCTCAATGCTCGCTCAGAGACTATTGATACC 840
Db 297 IleTyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThr 316
QY 841 AAAAACHAATCTCAGAAACCCAGACTCTGGGGCGTTCGAGCCAGTTCCTGTGGCCCTGC 900
Db 317 LysThrAsnCysArgAsnProAspCysTrpGlyValArgGlyGlnPheCysGlyProCys 336
QY 901 CTTCCGAAACCGTTATCGTGAAGAGTCAGGATGCTCTGTGATCCGAACTGGCATTGC 960
Db 337 LeuArgAsnArgTyrGlyGluGluValArgAspAlaLeuLeuAspProAsnTrpHisCys 356
QY 961 CCGCTTGTTCGAGGAATCTGCAACTGCACTGCTGCGCGAGCAGAGATGACCGTGTGCG 1020
Db 357 ProProCysArgGlyIleCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAla 376
QY 1021 ACTGGGCTCTGTGTATTGACCAATATCATGCTTGGGAATGTGCATGCTACTTG 1080
Db 377 ThrGlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeu 396
QY 1081 AAAAGCCTGAAACAGGAATTTGAAATGCAAGCA 1113
Db 397 LysSerLeuLysGlnGluPheGluMetGlnAla 407
```

RESULT 10

US-10-106-698-5654

; Sequence 5654, Application US/10106698

; Publication No. US20030109690A1

; GENERAL INFORMATION:

; APPLICANT: Ruben et al.

; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides

; FILE REFERENCE: PA005P1

; CURRENT APPLICATION NUMBER: US/10/106,698

```

; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: Patent In Ver. 3.0
; SEQ ID NO 5654
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-5654

Alignment Scores:
Pred. No.: 7,13e-177 Length: 407
Score: 1929.00 Matches: 364
Percent Similarity: 98.11% Conservative: 0
Best Local Similarity: 98.11% Mismatches: 7
Query Match: 98.32% Indels: 0
DB: 14 Gaps: 0

US-10-046-935-2234 (1-1116) x US-10-106-698-5654 (1-407)

QY 1 ATGACGCTCCCGCTGCCGAGAGATCTCAGAGTAAAGAGAACTTAAAGAAATTC 60
DB 37 MetAspAlaArgArgValProGlnLysAspLeuArgValLysLysAsnLeuLysLysPhe 56
QY 61 AGATATCTGAGTTGATTTCATCGAAACCTCGTCATCTCTGATGACAGTTGTGACAGC 120
DB 57 ArgTyrValLysLeuIleSerMetGluThrSerSerSerSerSerSerSerSerSerSer 76
QY 121 TTTCTCTCTGATAATTTTGCACACAGAGCTGCAGTCAGTTCCGAGGAGCTGTAGNCC 180
DB 77 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr 96
QY 181 CGCAGCAGTGCAGGCACTCTGACCTCTCAGGCTGGCGATGAAGTTTCCAGCGCCGAGT 240
DB 97 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 116
QY 241 ACCAGGGAGACACCAACAAAGACAGAGTCCGCCAGCCCTCAGAGAAATTCGTGACT 300
DB 117 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 136
QY 301 GATTCCAACTCCGATTCAGAGATGAAGTGAATGAATTTTGGAGAAAGGCTTTA 360
DB 137 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 156
QY 361 AATATAAGCAAAACAAAGCAATGCTTGCAGAACTCATGTCTGAATTAGAAAGCTTCCCT 420
DB 157 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 176
QY 421 GCGTCGTTCGTGGAGACATCCCTCCAGGCTCCGACTCACAATCAAGAGACCGCGA 480
DB 177 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 196
QY 481 AGCGGTACATTCCTCGGCTGTGTCTTCAGAGAGAAACCTCAACGAGAGCTCGCTCTT 540
DB 197 ArgArgThrPheProGlyValAlaSerArgArgAsnProGluArgArgAlaArgProLeu 216
QY 541 ACCAGGTCAAGTCCCGGATCTCTCGGCTCCCTTGAAGCTCTACCCATGGRGAGCCAGT 600
DB 217 ThrArgSerArgSerArgIleLeuLysSerLeuAspAlaLeuProMetGluGluGlu 236
QY 601 TTKGAKMTWGTACATGTTGGTGGAGAAAGAGAGACCGGTGATGCTACATGAATGAA 660
DB 237 GluGluAspLysTyrMetLeuValArgLysThrValAspGlyTyrMetAsnGlu 256
QY 661 GATGACCTGCCAGAGCCGCTCCGATCATCGTGAACCTTCCGATATTAATTCG 720
DB 257 AspAspLeuProArgSerArgArgSerValThrLeuProHisIleIleArg 276

; Sequence 14, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MEM01-029P2RNM
; CURRENT APPLICATION NUMBER: US/10301822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 450
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-301-822-14

Alignment Scores:
Pred. No.: 4,43e-172 Length: 450
Score: 1879.50 Matches: 364
Percent Similarity: 80.89% Conservative: 0
Best Local Similarity: 80.89% Mismatches: 7
Query Match: 95.80% Indels: 79
DB: 14 Gaps: 1

US-10-046-935-2234 (1-1116) x US-10-301-822-14 (1-450)

QY 1 ATGAGCGCTCGCGGTGCCGAGAGATCTCAGAGTAAAGAGAACTTAAAGAAATTC 60

```

Db 1 MetAspAlaArgArgValProGlnLysAspLeuArgValLysAsnLeuLysPhe 20
QY 61 AGATATGCAAGTTGATTTCATGAACCTCGTCTCATCTCGATGACAGTTGTGACAGC 120
Db 21 ArgTyrValLysLeuLysMetGluThrSerSerSerAspSerCysAspSer 40
QY 121 TTTGCTTCTGATAATTTTCAACACG----- 147
Db 41 PheAlaSerAspAsnPheAlaAsnThrLysProLysPheArgSerAspLysSerGluGlu 60
QY 147 ----- 147
Db 61 LeuAlaSerValPheTyrGluAspSerAsnGluSerPheCysGlyPheSerGluSer 80
QY 147 ----- 147
Db 81 GluValGlnAspValLeuAspHisCysGlyPheLeuGlnLysProArgProAspValThr 100
QY 147 ----- 147
Db 101 AsnGluLeuAlaGlyLePheHisAlaAspSerAspAspGluSerPheCysGlyPheSer 120
QY 148 -----AGCTGCAOTCAGTTCCGGAAGCTGTAGACCCGC 193
Db 121 GluSerGluIleGlnAspGlyMetArgLeuGlnSerValArgGluGlyCysArgThrArg 140
QY 184 AGCAGTGCAGGCACCTCTGCACCTCTCAGGTCGCGATGAAGTTTCCAGCGCGAGTACC 243
Db 141 SerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSerThr 160
QY 244 AGGGAGCAACCAACAAAAAGCAGAGTCCCGCAGCCCTCAGAGAATTCGTGACTGAT 303
Db 161 ArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThrAsp 180
QY 304 TCCAATCCGATTGAGAGATGAAGTGAATGAATTTTTCAGAAAAAGGCTTTAAAT 363
Db 181 SerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeuAsn 200
QY 364 ATAAGCAAAACAAAGCAAGCTTGCAGAACTCATGTCTGAATTAGAAAGCTTCCCTGGC 423
Db 201 IleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPheProGly 220
QY 424 TCGTTCCTCGTGAAGACATCCCTCCAGGCTCCGACTCACAAATCAAGGAGACCGCGAAG 483
Db 221 SerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArgArg 240
QY 484 CGTACATTCGGGTGTGCTTCCAGAGAAACCTGAAACGAGAGCTCGTCTCTTACC 543
Db 241 ArgThrPheProGlyValAlaSerArgArgAsnProGluArgArgAlaArgProLeuThr 260
QY 544 AGGTCAAGGTCCGGATCCCTCGGTCCCTGACGCTCTACCCATGCGRGAGCCAGATTTK 603
Db 261 ArgSerArgSerArgIleLeuLysSerLeuAspAlaLeuProMetGluGluGluGlu 280
QY 604 GARMTWMTGATGTTGTTGAGAAAGAGGAGGACCGGTGGATGCTCATGAATGAAGAT 663
Db 281 GluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGluAsp 300
QY 664 GACCTGCCAGAACCGTCTCAGATCATCCGTGACCCCTCCGATATTAATTCGCCCA 723
Db 301 AspLeuProArgSerArgArgSerSerValThrLeuProHisIleLeuArgPro 320
QY 724 GTGGAAGAAATACAGAGGAGGAGTTGGAGAACGTCTGCAGCAATTTCTCGAGAGAAGATA 783
Db 321 ValGluGluIleThrGluGluGluLeuGluAsnValCysSerAsnSerArgGluLysIle 340
QY 784 TATACCGTTCACTGGGCTCTACTTGTTCATCAATGCCGTCCAGAGACTATTGATACCAA 843
Db 341 TyrAsnArgSerLeuGlySerThrCysHisGlnCysArgGlnLysThrIleAspThrLys 360
QY 844 ACAATCTCAGAAACCCAGACTGCTGGGCGCTTCGAGCCAGTCTGTGGCCCTTCCCT 903
Db 361 ThrAsnCysArgAsnProAspCysTrpGlyValArgGlyGlnPheCysGlyProCysLeu 380

QY 904 CGAAACCGTTATGTTGAAGAGTCCAGGATGCTCTCTCGGATCCGAACCTGGCATTCGCCG 963
Db 381 ArgAsnArgTyrGlyGluValArgAspAlaLeuLeuAspProAsnTrpHisCysPro 400
QY 964 CTTTGTCCAGGAATCTGCAACTGCAAGTTCTCCCGCAGCAGAGTGCAGGTGTGCGACT 1023
Db 401 ProCysArgGlyLeCysAsnCysSerPheCysArgGlnArgAspGlyArgCysAlaThr 420
QY 1024 GGGGCTCTGTCTATTAGCCAAATATCATGGCTTTGGGAATGTGCATCCCTACTTGA 1083
Db 421 GlyValLeuValTyrLeuAlaLysTyrHisGlyPheGlyAsnValHisAlaTyrLeuLys 440
QY 1084 AGCCTGAAACAGGAATTTCAAATGCAAGCA 1113
Db 441 SerLeuLysGlnGluPheGluMetGlnAla 450

RESULT 12

US-09-851-588-4
; Sequence 4, Application US/09851588
; Patent No. US20020042067A1
; GENERAL INFORMATION:

; APPLICANT: Mack, David
; APPLICANT: Gish, Kurt C.
; APPLICANT: Wilson, Keith E.
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSING COLORECTAL CANCER, COMPOSITIONS, AND
; TITLE OF INVENTION: OF SCREENING FOR COLORECTAL CANCER MODULATORS
; FILE REFERENCE: A-68829-1/DJB/JJD/AMS
; CURRENT APPLICATION NUMBER: US/09/851,588
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 09/642,252
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: US 09/656,002
; PRIOR FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 4
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-851-588-4

Alignment Scores:
Pred. No.: 1.48e-150 Length: 347
Score: 1656.00 Matches: 338
Percent Similarity: 97.41% Conservative: 0
Best Local Similarity: 97.41% Mismatches: 9
Query Match: 84.40% Indels: 2
DB: Gaps: 0

US-10-046-935-2234 (1-1116) x US-09-851-588-4 (1-347)

QY 1 ATGACGCTCGCCGTCGCGCAGAAAGATCTCAGAGTAAAGAAAGAACTTAAAGAAATTC 60
Db 1 MetAspAlaArgArgValProGlnLysAspLeuArgValLysAsnLeuLysPhe 20
QY 61 AGATATGCAAGTTGATTTCATGAACCTCGTCTCATCTCGATGACAGTTGTGACAGC 120
Db 21 ArgTyrValLysLeuLysMetGluThrSerSerSerAspSerCysAspSer 40
QY 121 TTTGCTTCTGATAATTTTCAACACGAGGCTGAGTCCGGAAGGCTGTAGGACC 180
Db 41 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr 60
QY 181 CGCAGCAGTGCAGGCACCTCTGGACCTCTCAGGGTGGCGATGAAGTTTCCAGCGCGAGT 240
Db 61 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 80
QY 241 ACCAGGGAGCAACCAACAAAAAGCAGAGTCCCGCAGCCCTCAGAGAAATCTGTGACT 300
Db 81 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 100
QY 301 GATTCACACTCCGATTCCAGAAAGATGAAGTGAATTTTGGAGAAAGGCGCTTTA 360

```
Db 101 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 120
Qy 361 AATATAAAGCAAAACAAAGCAATGCTTCCAAACATCATGTCGAATTAAGAGCTCCCT 420
Db 121 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140
Qy 421 GGCTCGTTCGTTGGAGACATCCCTCCAGGCTCCGACTCACAATCAAGGAGACCGGA 480
Db 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 160
Qy 481 AGGGTACATTCGCGGTGTTGCTTCCAGAGAAACCTGAACGAGAGCTCGTCTCTT 540
Db 161 ArgArgThrPheProGlyValAlaSerArgAsnProGluArgAlaArgProLeu 180
Qy 541 ACCAGGTCAAGCTCCGAGTCTCGGGTCCCTTGACGCTCTACCCATGGRGAGCCAGT 600
Db 181 ThrArgSerArgSerArgGlyLeuGlySerLeuAspAlaLeuProMetGluGluGlu 200
Qy 601 TTGARKMTMGATACATGTTGGTGAGAAAGAGAACCGTGGATGGCTACATGAATGAA 660
Db 201 GluGluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 220
Qy 661 GATGACTGCCCAAGCCGCTCGATCCAGATCATCCGTGACCTTCCGCATATATTCGC 720
Db 221 AspAspLeuProArgSerArgArgSerArgSerValThrLeuProHisIleLeuArg 240
Qy 721 CCAGTGAAGAAATACAGA-GGAGGAGTTGGAGAAAGAGAACCGTGGATGGCTACATGAATGAA 778
Db 241 ProValGluGluLeuThrGluGlyGlyValGluArgLeuGlnPheSerLysArg 260
Qy 779 AGATATATAACCGTCACTGGCTCTACTTGTCTCAATGCGTCAAGAGACTATTGATA 838
Db 261 ArgTyrIleThrValHisTyrAlaLeuLeuValIleAsnAlaValArgArgLeuLeuLe 280
Qy 839 CCAAAACAACTGCAGAAACCCAGACTGCTGGGGCGTTGAGGCCAGTTCGTGGCCCT 898
Db 281 ProLysGlnThrAlaGluThrGlnThrAlaGlyAlaPheGluAlaSerSerValAlaPro 300
Qy 899 GCCTTCCAAACCGTATGTTGTAAGAGTTCAGGATGCTCTGCTGGATCCGAACTGGATT 958
Db 301 AlaPheGluThrValMetValLysArgSerGlyMetLeuGlyTyrIleArgThrGlyLe 320
Qy 959 GCGCGCTTGTGAGGAATCTGCAACTGCACTGCTGGCGGCTTCCGCGCAGAGATGGAGGTG 1018
Db 321 AlaArgLeuValGluGluSerAlaThrAlaValSerAlaGlySerGluMetAspGlyVal 340
Qy 1019 CGACTGGGTCCTGTGTATT 1039
Db 341 ArgLeuGlySerLeuCysIle 347
```

RESULT 13

```
US-10-295-027-1194
; Sequence 1194, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter H.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
```

```
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1194
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-1194

Alignment Scores:
Pred. No.: 1,48e-150 Length: 347
Score: 1656.00 Matches: 338
Percent Similarity: 97.41% Conservative: 0
Best Local Similarity: 97.41% Mismatches: 9
Query Match: 84.40% Indels: 2
DB: 14 Gaps: 0

US-10-046-935-2234 (1-1116) x US-10-295-027-1194 (1-347)
Qy 1 ATGACGCTCGCCGCTGCCGCGAAGAGATCTCAGAGTAAGAGAACTTAAGAAATTC 60
Db 1 MetAspAlaArgArgValProGlnLysAspLeuArgValLysLysAsnLeuLysPhe 20
Qy 61 AGATATGTGAAGTTGATTTCATGGAAACCTCGTCATCCTCTGATGACAGTGTGACAGC 120
Db 21 ArgTyrValLysLeuLeuSerMetGluThrSerSerSerSerSerSerSerCysAspSer 40
Qy 121 TTTGCTTCTGATATTTTGCACACACAGAGCTCGAGTTCGGGAAGGCTGTAGGACC 180
Db 41 PheAlaSerAspAsnPheAlaAsnThrArgLeuGlnSerValArgGluGlyCysArgThr 60
Qy 181 CGCAGCCAGTGCAGGCACTCTGGACCTCTCAGGGTGGCGATGAAGTTTCAGCGCGAGT 240
Db 61 ArgSerGlnCysArgHisSerGlyProLeuArgValAlaMetLysPheProAlaArgSer 80
Qy 241 ACCAGGGAGCAACCAACAAAAAGCAGAGTCCCGCCAGCCCTCAGAGATTCGTGACT 300
Db 81 ThrArgGlyAlaThrAsnLysLysAlaGluSerArgGlnProSerGluAsnSerValThr 100
Qy 301 GATTCACATCCGATTTCAGAGATGAAGTGGAAATGAATTTTGGAGAAAAGGGCTTTA 360
Db 101 AspSerAsnSerAspSerGluAspGluSerGlyMetAsnPheLeuGluLysArgAlaLeu 120
Qy 361 AATATAAAGCAAAACAAAGCAATGCTTCCAAACATCATGTCGAATTAAGAGCTTCCT 420
Db 121 AsnIleLysGlnAsnLysAlaMetLeuAlaLysLeuMetSerGluLeuGluSerPhePro 140
Qy 421 GGCTCGTTCGTTGGAGACATCCCTCCAGGCTCCGACTCACAATCAAGGAGACCGGA 480
Db 141 GlySerPheArgGlyArgHisProLeuProGlySerAspSerGlnSerArgArgProArg 160
Qy 481 AGGGTACATTCGCGGTGTTGCTTCCAGAGAAACCTGAACGAGAGCTCGTCTCTT 540
Db 161 ArgArgThrPheProGlyValAlaSerArgAsnProGluArgAlaArgProLeu 180
Qy 541 ACCAGGTCAAGCTCCGAGTCTCGGGTCCCTTGACGCTCTACCCATGGRGAGCCAGT 600
Db 181 ThrArgSerArgSerArgGlyLeuGlySerLeuAspAlaLeuProMetGluGluGlu 200
Qy 601 TTGARKMTMGATACATGTTGGTGAGAAAGAGAACCGTGGATGGCTACATGAATGAA 660
Db 201 GluGluAspLysTyrMetLeuValArgLysArgLysThrValAspGlyTyrMetAsnGlu 220
Qy 661 GATGACTGCCCAAGCCGCTCGATCCAGATCATCCGTGACCTTCCGCATATATTCGC 720
Db 221 AspAspLeuProArgSerArgArgSerArgSerValThrLeuProHisIleLeuArg 240
Qy 721 CCAGTGAAGAAATACAGA-GGAGGAGTTGGAGAAAGAGAACCGTGGATGGCTACATGAATGAA 778
Db 241 ProValGluGluLeuThrGluGlyGlyValGluArgLeuGlnPheSerLysArg 260
Qy 779 AGATATATAACCGTCACTGGCTCTACTTGTCTCAATGCGTCAAGAGACTATTGATA 838
Db 261 ArgTyrIleThrValHisTyrAlaLeuLeuValIleAsnAlaValArgArgLeuLeuLe 280
Qy 839 CCAAAACAACTGCAGAAACCCAGACTGCTGGGGCGTTGAGGCCAGTTCGTGGCCCT 898
Db 281 ProLysGlnThrAlaGluThrGlnThrAlaGlyAlaPheGluAlaSerSerValAlaPro 300
Qy 899 GCCTTCCAAACCGTATGTTGTAAGAGTTCAGGATGCTCTGCTGGATCCGAACTGGATT 958
Db 301 AlaPheGluThrValMetValLysArgSerGlyMetLeuGlyTyrIleArgThrGlyLe 320
Qy 959 GCGCGCTTGTGAGGAATCTGCAACTGCACTGCTGGCGGCTTCCGCGCAGAGATGGAGGTG 1018
Db 321 AlaArgLeuValGluGluSerAlaThrAlaValSerAlaGlySerGluMetAspGlyVal 340
Qy 1019 CGACTGGGTCCTGTGTATT 1039
Db 341 ArgLeuGlySerLeuCysIle 347
```


82	ATGAAACCTCGTCACTCTCTGATGACAGTTGTGACAGCTTTGCTTCTGATAATTTTGCA	141
1	MetGluThrSerSerSerAspSerCysAspSerPheAlaSerAspAsnPheAla	20
142	AACACGAGGTGCAGTTCGGGAGGCTGTAGACCCGCGAGCCAGTGCAGGCCTCT	201
21	AsnThrArgLeuGlnSerValArgGluGlyCysArgThrArgSerGlnCysArgHisSer	40
202	GGACCTCTCAGGGTGGCGATCAAGATTTCCAGCGCGAGTACCAGGGGAGCAACCAACA	261
41	GlyProLeuArgValAlaMetLysPheProAlaArgSerThrArgGlyAlaThrAsnLys	60
262	AAAGCAGAGTCCCGCCAGCCCTCAGAGAACTCTGTGACTATTCCAACTCCGATTCAGAA	321
61	LysAlaGluSerArgGlnProSerGluAsnSerValThrAspSerAsnSerAspSerGlu	80
322	GATCAAAAGTGGAAATGATTTTTTCGAGAAAAGGCTTTAAATAAAGCAACAACAAAGCA	381
81	AspGluSerGlyMetAsnPheLeuGluLysArgAlaLeuAsnLleLysGlnAsnLysAla	100
382	ATGCTTGCAAACTCATGTCTGAATTAGAAAGCTTCCCTCGCTCGTTCGGTGGAGACAT	441
101	MetLeuAlaLysLeuMetSerGluLeuGluSerPheProGlySerPheArgGlyArgHis	120
442	CCCTCTCCAGGCTCCGACTCACAACTCAAGAGAGCCGGAAGGCTACATCCCGGGTGT	501
121	ProLeuProGlySerAspSerGlnSerArgArgProArgArgThrPheProGlyVal	140
502	GCTTCCAGGAGAAACCTGAACGGAGAGCTCGTCTCTTTACCAAGGTCAGGTCCTCCGGATC	561
141	AlaSerArgArgAsnProGluArgArgAlaArgProLeuThrArgSerArgSerArgIle	160
562	CTCGGTCCTTCACGCTCTACCCATGGRGAGAGCCAGTTTTCAGRWTTWGTACATGTTG	621

Blank

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 12, 2004, 15:07:46 ; Search time 23 Seconds
(without alignments)
6435.729 Million cell updates/sec

Title: US-10-046-935-2234
Perfect score: 1962
Sequence: 1 atggacgctgcgcgctgcc.....aattgaaatgcaagcataa 1116

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 956278

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2_1/USPRO.spool_p/US10046935/runat_12102004.155226.15267/app_query.fasta_1.1287
-DB=Issued_Patents_AA -QPM=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdd
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10046935 @CGN 1.1 37 @runat_12102004.155226.15267 -NCPU=6 -ICPU=3
-NO MMAP -LARGEOQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREAS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pcp:*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pcp:*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pcp:*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pcp:*
5: /cgn2_6/ptodata/1/iaa/6C.COMB.pcp:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1656	84.4	347	US-09-851-588-4	Sequence 4, Appli
2	1520.5	77.5	320	US-09-851-588-2	Sequence 2, Appli
3	125.5	6.4	433	US-09-252-991A-20376	Sequence 20376, A
4	125	6.4	754	US-09-976-594-375	Sequence 375, App
5	116.5	5.9	904	US-09-976-594-615	Sequence 615, App
6	110	5.6	536	US-09-188-930-185	Sequence 185, App
7	110	5.6	536	US-09-312-283C-185	Sequence 185, App
8	110	5.6	742	US-08-791-115B-5	Sequence 5, Appli
9	107	5.5	380	US-09-270-767-43626	Sequence 43626, A
10	106	5.3	582	US-09-252-991A-21510	Sequence 21510, A
11	106	5.4	820	US-09-248-796A-17231	Sequence 17231, A
12	105	5.4	308	US-09-252-991A-31991	Sequence 31991, A

13	105	5.4	679	4	US-09-252-991A-27111	Sequence 27111, A
14	103.5	5.3	947	4	US-09-418-780A-1	Sequence 1, Appli
15	103.5	5.3	947	4	US-09-392-714-23	Sequence 23, Appli
16	102.5	5.2	499	4	US-09-252-991A-30259	Sequence 30259, A
17	102	5.2	266	4	US-09-252-991A-32512	Sequence 32512, A
18	102	5.2	301	4	US-09-248-796A-18853	Sequence 18853, A
19	102	5.2	310	4	US-09-270-767-42373	Sequence 42373, A
20	102	5.2	395	4	US-09-252-991A-30497	Sequence 30497, A
21	101.5	5.2	944	3	US-09-419-285A-2	Sequence 2, Appli
22	101.5	5.2	1850	4	US-09-620-093A-5	Sequence 5, Appli
23	100.5	5.1	3075	2	US-08-460-309-5	Sequence 5, Appli
24	100.5	5.1	3075	2	US-08-125-077-5	Sequence 5, Appli
25	100	5.1	559	1	US-08-320-559-31	Sequence 31, Appli
26	100	5.1	559	3	US-08-545-860D-31	Sequence 31, Appli
27	100	5.1	559	4	US-09-538-092-1261	Sequence 1261, Ap
28	100	5.1	559	5	PCT-US94-04496-31	Sequence 31, Appli
29	99	5.0	667	4	US-09-248-796A-19663	Sequence 19663, A
30	99	5.0	2200	4	US-09-796-575-2	Sequence 2, Appli
31	98.5	5.0	295	3	US-09-199-637A-341	Sequence 341, App
32	97.5	5.0	357	4	US-09-252-991A-20112	Sequence 20112, A
33	97.5	5.0	454	2	US-09-159-385-1	Sequence 1, Appli
34	97.5	5.0	454	3	US-09-186-277-1	Sequence 1, Appli
35	97.5	5.0	590	4	US-09-312-283C-409	Sequence 409, App
36	97.5	5.0	786	4	US-09-509-802-2	Sequence 2, Appli
37	97.5	5.0	787	3	US-09-188-930-334	Sequence 334, App
38	97.5	5.0	787	4	US-09-312-283C-334	Sequence 334, App
39	97.5	5.0	1342	2	US-08-484-438-9	Sequence 9, Appli
40	97.5	5.0	1523	4	US-09-538-092-955	Sequence 955, App
41	97	4.9	2802	3	US-09-542-331-1	Sequence 1, Appli
42	97	4.9	2802	3	US-09-510-791-1	Sequence 1, Appli
43	96.5	4.9	439	4	US-09-252-991A-21361	Sequence 21361, A
44	96.5	4.9	1953	4	US-09-917-254-92	Sequence 92, Appl
45	96	4.9	494	4	US-09-538-092-1282	Sequence 1282, Ap

ALIGNMENTS

RESULT 1
; Sequence 4, Application US/09851588
; Patent No. 6682890
; GENERAL INFORMATION:
; APPLICANT: Mack, David
; APPLICANT: Gish, Kurt C.
; APPLICANT: Wilson, Keith E.
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSING COLORECTAL CANCER, COMPOSITIONS, AND N
; FILE REFERENCE: A-68829-1/DJB/JJD/AMS
; CURRENT APPLICATION NUMBER: US/09/851,588
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 09/642,252
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: US 09/656,002
; PRIOR FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-851-588-4

Alignment Scores:
Pred. No.: 9.2e-177 Length: 347
Score: 1656.00 Matches: 338
Percent Similarity: 97.41% Conservative: 0
Best Local Similarity: 97.41% Mismatches: 9
Query Match: 84.40% Indels: 2
DB: 4 Gaps: 0

US-10-046-935-2234 (1-1116) x US-09-851-588-4 (1-347)

QY 1 ATGGACGCTGCGCGCTGCGCAGAAAGATCTCAGAGTAAGAAGACTTAAAGAAATTC 60

[illegible]

```

; Sequence 2, Application US/09851588
; Patent No. 682890
; GENERAL INFORMATION:
; APPLICANT: Mack, David
; APPLICANT: Gish, Kurt C.
; APPLICANT: Wilson, Keith E.
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSING COLORECTAL CANCER, COMPOSITIONS AND METHODS OF SCREENING FOR COLORECTAL CANCER MODULATORS
; TITLE OF INVENTION: OF SCREENING FOR COLORECTAL CANCER MODULATORS
; FILE REFERENCE: A-68829-1/DJB/JJD/AMS
; CURRENT APPLICATION NUMBER: US/09/851,588
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 09/642,252
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: US 09/656,002
; PRIOR FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-851-588-2

Alignment Scores:
Pred. No.: 1,44e-161 Length: 320
Score: 1520.50 Matches: 306
Percent Similarity: 96.26% Conservative: 3
Best Local Similarity: 95.33% Mismatches: 10
Query Match: 77.50% Indels: 2
DB: 4 Gaps: 1

```



```
DB: 4 Gaps: 12
US-10-046-935-2234 (1-1116) x US-09-976-594-375 (1-754)
QY 4 GAGCTCGCGGTGCGGAGAAAGATCTCAGAGTAAAGAACTTAAAGAAATTCAGA 63
Db 437 AspIleArgAsnSerGluLysGluAspLysTyrLysAsnLysValLysLysArgAla 456
QY 64 TATGTGAAGTTGATTTTCATGGAACCTCGTCTCTCTCATGACAGTTGTGACAGCTTT 123
Db 457 LysSerLysSerArgSerLysLysGluLysSerLysSerLysGluLysArgAspSerLys 476
QY 124 GTCTCTGATTAATTTGCAACACAGAGCTGCAGTTCGGGAAGCTGTAGACCCGC 183
Db 477 HisAsnArgAsn-----GluGluLysArgMetArg 486
QY 184 AGCAGTGCAGGCACTCTGACCTCTCAGGTTGGCGATGAAGTTTCAGCGCGGAGTACC 243
Db 487 SerArgSerLysGlyArgAspHisGluAsnValLysGluLysGluLysGlnSerAspSer 506
QY 244 AGGGAGCAACCAACAAAGAGCAGTCCCGCAGCCCTCAGAGAAATTCGTGACTGAT 303
Db 507 LysGlyLysAspGlnGluArgSerArgSerLysGluLysSerLysGln-----LeuGlu 524
QY 304 TCCAACCTCCGATTCAGAAGATGAAGTGAATGAATTTTGGAGAAAAGGCTTTAAAT 363
Db 525 SerLysSerAsnGluHisAspHisSerLysSerLysGluLysAspArgArgAla----- 542
QY 364 ATAAACCAACAAACCAATGCTTCGAAACTCATGTCGATTAATAGAAAGCTTCCCTGGC 423
Db 543 -----GlnSerArgSerArgGluCysAspIleThrLysGlyLysHisSerTyrAsnSer 560
QY 424 TCGTTCCTCGTGGAGACATCCCTCCAGGCTCCGACTCACAATCAAGAGACCGCGAAGG 483
Db 561 ArgThrArgGluArg-----SerArgSerArgAspArgSerArgValArgSer 577
QY 484 CGTACATTCGCGGTGTGCTTCAGG-----AGAAACCTCGAA 522
Db 578 ArgThrHisAspArgAspArgSerArgSerLysGluLysArgTyrArgGluGlnGlu 597
QY 523 CGGAGAGCTCGTCTCTTACAGTCAAGTCCGAGTCCGATCTCGGTCCTTCCTGAGCTCTA 582
Db 598 TyrArgArgArgGlyArgSerArgSerArgGluArg-----ArgThr 611
QY 583 CCCATGGRGAGACCGATTTKGARKMTWGTATCATGTTGGTGAAGAGAGAGACCGTG 642
Db 612 ProGlyArgSerArgSerLysAsp-----ArgArgArgArgArgArg 626
QY 643 GATGGCTACATGAATGAGATGACCTCCAGAGACCGTCTGCTCCAGATCATCGTGACC 702
Db 627 AspSerArgSerGluArgGluGluSerGlnSerArgAsnLysAspLys----- 643
QY 703 CTTCCGCATATAATTCGCCAGTGGAGAAATACAGAGGAGGAGTGGAGAGCTGTCG 762
Db 644 -----TyrArgAsnGlnGluSerLysSerHisArgLysGluAsn----- 657
QY 763 AGCAATTCGAGAGATATATAACCGTTCACTGGGC---TCTACTTGTCTCAATGTC 819
Db 658 SerGluSerGluLysArgMetTyrSerLysSerArgAspHisAsnSerAsnSer 677
QY 820 CGTCAGAGACTATTGAT-----ACCAAAACAACACTGCAGAAACCCA 861
Db 678 ArgGluLysLysAlaAspArgAspGlnSerProPheSerLysLysLysGlnSerGln 697
QY 862 GACTGCTGGGCGGTTCGAGGCGAGTCTGTGGCCCTCGCTTGGAAACCGTTATGGTGA 921
Db 698 Asp-----AspGluLysSerSerMetLeuLysAsnLysGluAspGlu 712
QY 922 GAGGTCCAGGATGCTCTGCTGGATCCGAAC 951
Db 713 LysIleArgSerValGluLysGluAsn 722
```

RESULT 5

```
US-09-976-594-615
; Sequence 615, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCES: PA-0041 US
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 615
; LENGTH: 904
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; FEATURE:
; OTHER INFORMATION: Incyte ID No. 6673549 2789525CD1
US-09-976-594-615
```

```
Alignment Scores:
Pred. No.: 0.00121 Length: 904
Score: 116.50 Matches: 65
Percent Similarity: 37.60% Conservative: 29
Best Local Similarity: 26.00% Mismatches: 111
Query Match: 5.94% Indels: 45
DB: 4 Gaps: 11
```

US-10-046-935-2234 (1-1116) x US-09-976-594-615 (1-904)

```
QY 16 GTCCGCGAGAAAGATCTCAGAGTAAAGAAAGAACTTAAAGAAATTCAGATATGTG----- 69
Db 119 IleProSerAlaPheLeuGluLeuLysGluGluLysGlnArgGlnIleGluGln 138
QY 70 ---AAGTTGATTTCCATGGAACCTCGTCACTCTCATCTCATGACAGATTTGTGACAGCTTTGCT 126
Db 139 GluLysLeuAlaSerMetLysLysGlnAspGluAspLysAspLysArgAspLysGluGlu 158
QY 127 TCTGATAATTTTCAAAACACGAGCTGCAGTCA-----GTTCCGGAAGCTGTAGG 177
Db 159 LysGluSerArgGluLysArgGluArgSerArgSerProArgArgLysSerArg 178
QY 178 ACCCGCAGCCAGTGCAGGCACTCTCGACCTCTCAGGTCGCGATGAAG---TTCCAGCG 234
Db 179 SerProSerProArgArgArgSerSerProValArgArgGluArgLysArgSerHisSer 198
QY 235 CGAGTACCGAGGAGCAACCAAAAGAGAGTCCCGCAGGCTCCAGAGAAATTCCT 294
Db 199 ArgSerProArgHisArgThrLysSerArgSerProSerProAlaProGluLysLysGlu 218
QY 295 GTCACTGATTCACCACTCCGATTCAGAGATGAAGTGAATGATTTTGGAGAAAAGG 354
Db 219 LysThrProGlu-----LeuProGluPro 226
QY 355 GCTTTAAATATAAGCAAAACAAAGCAATGCTTGCAAAACTCATGTCTGAATTA----- 408
Db 227 SerValLysValLysGluProSerValGlnGluAlaThrSerThrSerAspIleLeuLys 246
QY 409 -----GAAAGCTTCCTCGGTCGTTCCGTGGAAGACATCCCTCCAGAGTCC 456
Db 247 ValProLysProGluProIlePro-----GluProLysGluProSerProGluLys 263
QY 457 GACTCTCAATCAAGGAGACCGCGAAGCGGTACATTCCTGGGTGTGCTGCCAG----- 510
Db 264 AsnSerLysLysGluLysGluLysArgProArgProArgSerArgSerArgSerLys 283
QY 511 -----AGAAACCTTGAACCGGAGAGCTCGTCTCTTACCAGGTCA 549
Db 284 SerArgSerArgThrArgSerArgSerProSer---HisThrArgProArgArgHis 302
```

```
QY 550 AGGTCCTCGGATCTCTGGGTCCTTACGCTCTACCCATGGRGAGACGAGTTCGARKMT 609
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
303 ArgSerArg-----SerArgSerTyrSerProArgArgProSerProArgArg 319
QY 610 MWGTACATGTTGGTGAGAAAGAGGAGACCGTGGATGGCTACATGAATGAAGATGACCTG 669
Db ::|||::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
320 ArgProSer-----ProArgArgThrProProArgArgMetProPro----- 335
QY 670 CCCAGAACCGTCGCTCCAGATCATCCGTG 699
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
336 ProArgHisArgArgSerArgSerProVal 345

RESULT 6
US-09-188-930-185
; Sequence 185, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188.930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 185
; LENGTH: 536
; TYPE: PRT
; ORGANISM: mouse
US-09-188-930-185

Alignment Scores:
Pred. No.: 0.00472 Length: 536
Score: 110.00 Matches: 61
Percent Similarity: 37.28% Conservative: 43
Best Local Similarity: 21.86% Mismatches: 111
Query Match: 5.61% Indels: 64
DB: 3 Gaps: 12

US-10-046-935-2234 (1-1116) x US-09-188-930-185 (1-536)
QY 172 TGTAGACCCGC-----ACCCAGTCGAGCAC 198
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
250 CysArgProArgProArgAlaCysAlaSerLeuLeuLeuLeuMetGlnArgCysTrpHis 269
QY 199 TCTGGACCT---CTCAGGGTGGCGATGAAGTTTCCAGCGCGAGTACCAGGGGAGCAACC 255
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
270 AlaAspProGlnValArgProThrPheGlnGluLeuThrSerGluThrGluAspLeuCys 289
QY 256 AACAAAAACGACAG-----TCCGCCAGCCCTCAGAGAATCTGTGACT 300
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
290 GluLeuProAspGluGluValLysAspLeuAlaHisGluProGlyGluLysSerSerLeu 309
QY 301 GATTCAACCTCCGATTCAGAAAGTGAAGTGAATTTTGGAGAAAGGCGTTTA 360
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
310 GluSerLysSerGluAlaArgProGluSerSerArg-----LeuLysArgAlaSerAla 327
QY 361 AATATAAGCAAAACGAATGCTTGCACAACTCATGTCTGAATTAGAAAGC----- 414
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
328 ProProPheAspAsnAspCysSerLeuSerGluLeuLeuSerGlnLeuAspSerGlyLe 347
QY 415 TTCCCTGGCTCGTTCGCTGGAAGACATCCCTCCAGGCTCCGACTCACAATCAAGGAGA 474
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
348 PheArgLeuLeuLysGlyProGluGluLeuSerArgSerSerGluCysLysLeu 367
QY 475 CCGCGAAGG-----CGTACATTCGCGGGTGTCTTCCAGGAGAAACCTGACGG 525
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
368 ProSerSerSerGlyLysArgLeuSerGlyValSerSerValAspSerAlaPheSer 387
```

```
QY 526 AGAGTCGTCTCTTACAGGTCAGGTCGCGATCTCGGGTCCCTTGACGCTCTACCC 585
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
388 SerArgGlySerLeuSerPheGluArgGluAlaSerThrGlyAspLeuGlyPro 407
QY 586 ATGGRGAGAGCCAGTTTKGARKMTMWGTACATGTTGGTGAGAAAGAGGAGACCGCTGAT 645
Db ::|||::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
408 ThrAsp-----lleGlnLysLysLysLeuValAsp 417
QY 646 GGCTACATGAATGAAGATGACCTGCCAGAGACCGCTGCTCCAGA----- 690
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
418 AlaIleIleSerGlyAsp-----ThrSerArgLeuMetLysIleLeuGlnProGln 434
QY 691 -----TCATCCGTGACCTTCGCGATATAAATTCGCCCGCTGCAA 729
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
435 AspValAspLeuValLeuAspSerSerAlaSerLeuLeuHisLeu-----AlaVal 451
QY 730 GAAATTACAGAGGAGGAGTTCGGAAC-----GTCTGCACGAATTCCTCGAGAGAATA 783
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
452 GluAlaGlyGlnGluCysValLysTrpLeuLeuLeuAsnAlaAsnProAsnLeu 471
QY 784 TATAACCGTTCACCTGGGCTCTACTTGCATCAATGCCGCTCAGAAG----- 828
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
472 ThrAsnArgLysGlySerThrProLeuHisMetAlaValGluArgLysGlyArgGlyLe 491
QY 829 -----ACTATTGATACCAAAACAACTGCAGAAACCCAGACTGCTGG 870
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
492 ValGluLeuLeuAlaArgLysThrSerValAsnAlaLysAspGluAspGlnTrp 510

RESULT 7
US-09-312-283C-185
; Sequence 185, Application US/09312283C
; Patent No. 6573095
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishnand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c2
; CURRENT APPLICATION NUMBER: US/09/312.283C
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 185
; LENGTH: 536
; TYPE: PRT
; ORGANISM: Mouse
US-09-312-283C-185

Alignment Scores:
Pred. No.: 0.00472 Length: 536
Score: 110.00 Matches: 61
Percent Similarity: 37.28% Conservative: 43
Best Local Similarity: 21.86% Mismatches: 111
Query Match: 5.61% Indels: 64
DB: 4 Gaps: 12

US-10-046-935-2234 (1-1116) x US-09-312-283C-185 (1-536)
QY 172 TGTAGACCCGC-----AGCCAGTCGAGCAC 198
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
250 CysArgProArgProArgAlaCysAlaSerLeuLeuLeuLeuMetGlnArgCysTrpHis 269
QY 199 TCTGGACCT---CTCAGGGTGGCGATGAAGTTTCCAGCGCGAGTACCAGGGGAGCAACC 255
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
270 AlaAspProGlnValArgProThrPheGlnGluLeuThrSerGluThrGluAspLeuCys 289
QY 256 AACAAAAACGACAG-----TCCGCCAGCCCTCAGAGAATCTGTGACT 300
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
290 GluLeuProAspGluGluValLysAspLeuAlaHisGluProGlyGluLysSerSerLeu 309
QY 301 GATTCAACCTCCGATTCAGAAAGTGAAGTGAATTTTGGAGAAAGGCGTTTA 360
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
310 GluSerLysSerGluAlaArgProGluSerSerArg-----LeuLysArgAlaSerAla 327
QY 361 AATATAAGCAAAACGAATGCTTGCACAACTCATGTCTGAATTAGAAAGC----- 414
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
328 ProProPheAspAsnAspCysSerLeuSerGluLeuLeuSerGlnLeuAspSerGlyLe 347
QY 415 TTCCCTGGCTCGTTCGCTGGAAGACATCCCTCCAGGCTCCGACTCACAATCAAGGAGA 474
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
348 PheArgLeuLeuLysGlyProGluGluLeuSerArgSerSerGluCysLysLeu 367
QY 475 CCGCGAAGG-----CGTACATTCGCGGGTGTCTTCCAGGAGAAACCTGACGG 525
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
368 ProSerSerSerGlyLysArgLeuSerGlyValSerSerValAspSerAlaPheSer 387
```

QY 301 GATTCCAACTCCGATTGACAGATGAAGTGAATGAAATTTTGGAGAAAAGGGCTTAA 360
Db 310 GluSerLysSerGluAlaArgProGluSerSerArg-----LeuLysArgAlaSerAla 327
QY 361 AATATAAGCAAAACAAACAACTGCTGCAAACTCATCTGTAATAGAAAGC----- 414
Db 328 ProProPheAspAsnAspCysSerLeuSerGluLeuSerGlnLeuAspSerGlyIle 347
QY 415 TTCCCTGGCTGCTTCGCTGGAAGACATCCCTCCAGGCTCCGACTCACAATCAAGGAGA 474
Db 348 PheProArgLeuLeuLysGlyProGluGluLeuSerSerSerGluCysLysLeu 367
QY 475 CCGCGAAGG-----CGTACATTCCTCCGCTGCTTCCAGAGAAACCTGAAACGG 525
Db 368 ProSerSerSerGlyLysArgLeuSerGlyValSerSerValAspSerAlaPheSer 387
QY 526 AGAGCTGCTCTTACCAGGTCAAGTCCCGGATCTCGGGTCCCTTGACGCTTACCC 585
Db 388 SerArgGlySerLeuSerLeuSerPheGluArgGluAlaSerThrGlyAspLeuGlyPro 407
QY 586 ATGGRGAGACCCAGTTTGARKMTWGTACATCTTGTGAGAAAGAGAGACCGTGGAT 645
Db 408 ThrAsp-----IleGlnLysLysLysLeuValAsp 417
QY 646 GGCTACATGAATGAAGATGACCTGCCAGAGCGCTCGCTCCAGA----- 690
Db 418 AlaIleSerGlyAsp-----ThrSerArgLeuMetLysIleLeuGlnProGln 434
QY 691 -----TCATCGTGACCTCCCGATATATATTCCTCCAGTGGAA 729
Db 435 AspValAspLeuValLeuAspSerSerAlaSerLeuLeuHisLeu-----AlaVal 451
QY 730 GAATTTACAGAGAGAGTTGGAGAAC-----GTCTGACGCAATCTCGAGAGAGATA 783
Db 452 GluAlaGlyGlnGluCysValLysTrpLeuLeuLeuAsnAlaAsnProAsnLeu 471
QY 784 TATAACCGTTCACTGGCTCTACTTGTCTCAATGCGCTCAGAG----- 828
Db 472 ThrAsnArgLysGlySerThrProLeuHisMetAlaValGluArgLysGlyArgGlyIle 491
QY 829 -----ACTATTGATACCAAAACAACTGCGAAGACCAGACTGCTGG 870
Db 492 ValGluLeuLeuAlaArgLysThrSerValAsnAlaLysAspGluAspGlnTrp 510

RESULT 8

US-08-791-115B-5
; Sequence 5, Application US/08791115B
; Patent No. 6262242
; GENERAL INFORMATION:
; APPLICANT: Steck, Peter
; APPLICANT: Perabhouse, Mark A.
; APPLICANT: Jasser, Samar
; APPLICANT: Yung, W.K. Alfred
; APPLICANT: Tavtigan, Sean V.
; TITLE OF INVENTION: A TUMOR SUPPRESSOR DESIGNATED TS10Q23.3
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.
; STREET: 555 Thirteenth Street, N.W., Suite 701-E
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 22204
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/791,115B
; FILING DATE: 30-JAN-1997
; CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
; NAME: Inhen, Jeffrey L.
; REGISTRATION NUMBER: 38,957
; REFERENCE/DOCKET NUMBER: 2318-134.A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-683-6040
; TELEFAX: 202-683-7031
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 742 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-791-115B-5

Alignment Scores:
Pred. No.: 0.00576 Length: 742
Score: 110.00 Matches: 79
Percent Similarity: 32.20% Conservative: 44
Best Local Similarity: 20.68% Mismatches: 164
Query Match: 5.61% Indels: 95
DB: 3 Gaps: 11

US-10-046-935-2234 (1-1116) x US-08-791-115B-5 (1-742)

QY 61 AGATATGTGAAGTTGATTTCCTGGAACCTCGTCTATCTCTGATGACAGTTGTGACAGC 120
Db 42 ArgThrLeuTyrrAlaLeuArgGlnAspThrArgSerAlaLeuGlyArgAspCysAlaGln 61
QY 121 TTTGCTCTCGATAATTTTGCAACACGAGCTCGAGTTCGGGAAGCTGTAGGACC 180
Db 62 PheSer-----Pro 64
QY 181 CGCAGCGCAGTGCAGGCACTCTGACACCTCTCAGGCTGCGATGAAGTTTCAGCGCGAGT 240
Db 65 LeuGlySerCysSerHisaspGlySerLeuArgValGluProLeuGlyGluAlaGlyLeu 84
QY 241 ACCAGGGAGCAACCAACAAAGAGAGTCCGCCAGCCCTCAGAGATTCTGTGACT 300
Db 85 ArgArgGlyArgGluThrAlaAlaAlaAlaAlaArgSerProSerGlnArgLeuAlaAla 104
QY 301 GATTCCAACTCCGATTGACAGATGAAGTGAATTTTGGAGAAAAGGGCTTAA 351
Db 105 AlaGlyAlaAlaProSerGlySerArgProAlaCysGlyGlySerGlyValSer 124
QY 352 AGGCGCTTTAAATATAAAGCAAAACAAAGCAATGCTTGCAAAACATCATGTCTGAATTAGAA 411
Db 125 ArgLeuLeuPheValPheSerAsnArgAlaAlaSerSerSerAlaSerProGluArgGlu 144
QY 412 AGCTTCCCTGCTGCTTCGCTGGAAGACATCCCTCCAGGCTCCGACTCACAATCAAG 471
Db 145 GlySerArgGly-----LeuGlyArgGluPro-----AlaGluAlaArg 157
QY 472 AGACCGGGAAGCGGTACATTCCCG-----GGT 498
Db 158 ArgArgArgArgArgHisLeuProLeuLeuGluArgGlyGlyGluAlaAlaAlaAla 177
QY 499 GTTGCTTCCAGGAGAAACCTCGAACCGAGAGACTCGT---CCTCTTACAGTCAAGTCC 555
Db 178 AlaAlaAlaAlaAlaProGlyArgGlySerGluSerProValThrIleSerArgAla 197
QY 556 CGGATCCTCGGTCCTTTCAGCT-----CTACCCATGGGAGAGCCAGTTTGAR 606
Db 198 GlyAsnAlaGlyGluLeuValSerProLeuLeuLeuProThrArgArg----- 214
QY 607 KMTWGTACATGTTGGTGAGAAAGAGAACCTGGATGCTACATGAATGAAGATGAC 666
Db 215 -----ArgArgArgHisIleGlnGlyProGlyProValLeuAsn 228
QY 667 CTGCCCAGAGCCCGTCTCGATCATCCCTGAGCCCTCCGATATAATTCGCCCATG 726
Db 229 LeuPro-----SerAlaAlaAlaAlaProProValAlaArgAlaPro 242

Db 295 IlePhePhe-----ProLeuSer 300
QY 790 GGTATATATCTTCTCGAGAATTCTCGAGAGCTTCTCAACTCTCTCTCTGTAATT 731
Db 301 GlyLysValValSerPheSerGly-:::|||||:::|LysSerProSerProCys- 314
QY 730 CTTCCACTGGCGAATATATGCGGAAGGCTCGAGATGATCTGGAGCGAGCGGCTTCGG 671
Db 315 -----GlySerArg----- 317
QY 670 GCAGGTCACTTCTCATTCATGATAGCCAT-----CCACGG 638
Db 318 SerSerHisValGluAlaPhePheHisArgSerProAlaAlaAspProHisCysProArg 337
QY 637 TCCTCTCTTCTCACCAACATGT-----ACWKAQMYTCWAACTGGCTCTCYCCATGG 584
Db 338 ProGlyAlaArgSerProValCysLeuGlyAla-ValAlaAspProGlyAlaAlaAlaG 357
QY 583 GTAGAGCGTCAAGGACCCGAGGATCCG-----GGACCTTGACCTGTGAAGAGGAC 533
Db 357 uGlnHisLeuProAlaProGluAspProValAlaLeuGly-----ProGlyArgArg 375
QY 532 GAGCTCT-----CGTTCAGGTTTCTCTGGAAGCAACACCCGGGATGTAC 485
Db 375 rGlyArgProGlyProLeuProAlaArgProGly-LeuProAlaThrGlyGluGluLea 395
QY 484 GCCTTCGCGGTCTCTGATTGTGAGTCGGAGCCTCGGAGGGGATGCTTCCACGGAACG 425
Db 395 rGSerArgGlyLeuAlaAspArgSerArgGlnThrGlyThrGlnAlaProGlyArgProG 415
QY 424 AGCCAGGGAAG 414
Db 415 luProGlyGlu 418

RESULT 11

US-09-248-796A-17231
; Sequence 17231, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09-248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR FILING DATE: 1998-02-13
; PRIOR FILING DATE: 1998-02-13
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17231
; LENGTH: 820
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-17231

Alignment Scores:

Pred. No.:	0.0172	Length:	820
Score:	106.00	Matches:	60
Percent Similarity:	34.09%	Conservative:	45
Best Local Similarity:	19.48%	Mismatches:	118
Query Match:	5.40%	Indels:	85
DB:	4	Gaps:	10

US-10-046-935-2234 (1-1116) x US-09-248-796A-17231 (1-820)

QY 75 GATTTCATGGAACCTCTCATCTCTGATGAC-----AGTTGTGACAGCTTGTCTTC 128
Db 200 AspSerHisGlyAsnIleThrIle-AspAspLysIleSerLysAsnGlyIleAlaThr 219
QY 129 TCATATATTTCGAACACCGAGCTCGAGTCAGTTCGGGAAGGCTGTAGGACCCGAGCCA 188
Db 219 rAsnAsnAsnThrThrThrProIleThrThrThrAlaAsnSerAsnAlaThrValGlyAs 239

QY 189 GTCAGGCACTCTGGACCTCTCAGGTGGCGATGAAGTTTCCAGCGCGGAGTACCAGGG 248
Db 239 nGlyAsnHisAsnThrProIleAsn-----AlaSerThrSerAsnG 253
QY 249 AGCAACCAACAAAAAGCAGAGTCCGCCAGCCCTCAGAG----- 288
Db 253 yIleThrLysLysThrLysCysLysLysAspSerThrLeuGluAlaLysIleAlaLe 273
QY 289 -AATTCTGTCACTGATTCCAACTCCGATTCCAGAAAGATGAAGTGAATGTTTGG 347
Db 273 uAsnSerAsnThrAspLeuLeuAsnGlnIleLysSerSerGluAsnThrThrValIleSe 293
QY 348 GAAAGAGCGCTTTAAATATAAGCAAAACAAAGCAATCTTGCAAACTCATGCTGAATT 407
Db 293 rProGluAspIleAsnLeuLysSerSerLysProThrProLeuLysSerLeuThr----- 311
QY 408 AGAAGCTTCCCTGGCTCGTCCGTGGAAGACATCCCTC-----CCAGCTCCGACTCACA 464
Db 312 -----ProAsnAsnAsnLysThrSerSerProIleValProIleSerSerSerGl 328
QY 465 ATCAAGGAGACCGCGAAGCGGTACATTCCCGGGTGTGCTTCCAGGAGAAACCTCGAACG 524
Db 328 nGlnGlnProProGlnGln-----GlnG 336
QY 525 GAGAGCTCTCTCTTACCAGGTCAAGTCCCGGATCCTCGGTCCCTTGACGCTCTACC 584
Db 336 nGluLysGlnProIleAlaCysThrLysSer----- 346
QY 585 CATGCGAGAGCCAGTTTGGARKMTMGATCATGTTGTGAGAAGAGGAGAGACCGTGA 644
Db 347 -----LysAlaLysAlaLysSerThrThrThrAlaArgLysSerLysThrAla-- 363
QY 645 TGGCTACATGAATCAAGATGACCTGCCGAGAACCGTCCCTCCAGATCATCCGTGACCC 704
Db 364 -----SerAlaLysGlyThrProLeuAlaGluVa 373
QY 705 TCCCATATATTCGCCAGTGGGAAGAAATTACAGAGAGGAGTGGAGAAC----- 756
Db 373 lProIleProIleProIleGluValAspProGluLysHisLeuSerAsnMetGlnMe 393
QY 757 -----GTCTGCAGCAATCTCTCGAGAGAGATATA 785
Db 393 tSerTyLeuSerLysLeuLysGluGlnGluValIleArgAsnTyrlleGluValLeuTh 413
QY 786 TAACCGT-----TCACCTGG 800
Db 413 rAsnGlnIleLysGluLeuSerLeuValGlnAsnGlyValIleThrPheAspAlaLeuLy 433
QY 801 CTCCTACTTGTCAATGCCGTCAGAGACTATTGATACCAAAACAAACTCGAGAAACCC 860
Db 433 sSerAsnValLysTyAsnAsnLysLysIleLeuThrSerThrThrThrSerThrGlyPr 453
QY 861 AGACTGCTGGCGCTTCGAGGC 882
Db 453 oGlyThrThrGlyAlaGlyGly 460

RESULT 12

US-09-252-991A-31991
; Sequence 31991, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142

```
; SEQ ID NO 31991
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31991

Alignment Scores:
Pred. No.: 0.0122 Length: 308
Score: 105.00 Matches: 69
Percent Similarity: 36.36% Conservative: 35
Best Local Similarity: 24.13% Mismatches: 118
Query Match: 5.35% Indels: 64
DB: 4 Gaps: 12

US-10-046-935-2234 (1-1116) x US-09-252-991A-31991 (1-308)
QY 13 CGCGTGGCGCAGAGACTCAGAGTAAGAGAACTTAAAGAAATTCAGATATGTGAAG 72
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
62 ArgValProSerAlaAlaSerIleArgProThrLeuArgTrpCysAlaProVallys 81
QY 73 TTGATTTCATGGAAACCTCGTCATCTCTGATGACAGTTGTGACAGCTTTCTCTGAT 132
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
82 AlaProPheSerTrpProAsnSerSerAspTrpIleArgCysSerGlyIleAlaProGln 101
QY 133 AATTTTGCACACAGAGCTGAGTCAGTTCCGGAAAGCTGTAGGACCCGCGACGATGC 192
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
102 LeuIleAlaThrAsnGlyLeu---SerAlaArgArgAspCysArgCysArgValArgAla 120
QY 193 AGGCACTCTGGACCTCTCAGGTCGCGATGAAGTTTCCAGCGCGAGTACCGAGGGAGCA 252
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
121 ThrSerLeuProValProLeuSerProArgIleSerThrGlyGluSerValGlyAla 140
QY 253 ACCAACAAGAGAGAGTCCCGCAGCCCTCAGAGAAT-----TCTGTGACTGAT 303
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
141 Ser-----PheAlaSerSerArgArgSerArgThrAlaArgLeuSerProSerSer 158
QY 304 TCCAACTCCGATTCAGAGATGAAGTGGATGAATTTTGGAGAAAGGCTTTA--- 360
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
159 SerCysSerProSerAlaSerProGlyThrValProArgProArgArgAlaMetPro 178
QY 361 -----AATATAAGCAAAACAAAGCAATGCTTCCA 390
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
179 LysAlaArgProArgValThrCysThrArgAlaMetSerAsnGlyLysValTrpLysSer 198
QY 391 AAACATCATGCTGAATTAGAAGCTTCCCTGGCTCGTTCCGTTGGAAGA----- 438
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
199 LysAsnHisSerGlnThrLysSer---ProThrProPheSerGlyLysValSerGlyLeu 217
QY 439 -----CATCCCTCCAGGCTCCGACTCACAATCAAGAGAGACCGGAGCGGT 486
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
218 SerThrAlaIleHisSerValPro-----LeuArgIleIleArgSerPhe 232
QY 487 ACATTCGCGGTGTGCTTCCAGAGAAACCTGNAAGGAGAGCTCGTCTTACCAGG 546
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
233 ThrValSerGly---ArgSerArgTrpAsnGlyArgArgSerArgProThrSerGln 251
QY 547 TCAAGTTCGCGGATCCTCGGTCCTTGCAC-----GCTTACCCATCGGAGAGCC 597
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
252 AlaArgSerArgAlaValSerSerAlaGlnLeuSerThrAlaGlnProAlaSerArgArg 271
QY 598 AGTTTGARKMTWGTACATTTGGTGAAGAGAGACCGGTGGATGGCTACATGAAT 657
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
272 ProGlyAsn-----ArgArgArgGlnSerSerHisGlySerThrThr 285
QY 658 GAAGATGACCTGCCAGAGACCGTGGCTCCAGATCATCCGTCGACCTTCCGCATATAAT 717
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
286 -----SerArgArgArgPhe----- 291
QY 718 CGCCAGTGGAGAAATATACAGAGGAGGATGTGAGAACGTCTCGACGAATTTCTGAGAG 777
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
292 -----AspAlaSerAsnMetCysAlaProLeuArgArg 302
QY 778 AAGATATATAACCGTTCA 795

; SEQ ID NO 27111
; LENGTH: 679
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27111

Alignment Scores:
Pred. No.: 0.0199 Length: 679
Score: 105.00 Matches: 70
Percent Similarity: 40.43% Conservative: 25
Best Local Similarity: 29.79% Mismatches: 80
Query Match: 5.35% Indels: 60
DB: 4 Gaps: 13

US-10-046-935-2234 (1-1116) x US-09-252-991A-27111 (1-679)
QY 415 TTCCTGGCTCGTTCGTTGGAAGACATCCCTCCAGGCTCCGACTCACAATCAAGAGA 474
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
323 PheProAlaSerSerArgTrpArgHisSerAlaProProArgArgProGlyAlaArgArg 342
QY 475 -----CCGCGAAGCGCT-----ACATTCCCGGGTGTGTTCTCCAGGAGA 513
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
343 AlaArgProSerProArgArgArgAspTrpAlaGlyProGlyAlaProAlaArgArg 362
QY 514 AACCTGAACGAGAGCTCGT-----CCTCTTACCAGTCAAGT----- 553
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
363 GlyArgGluArgArgAlaArgAlaGlyLeuProAlaIle-GlyAlaGlyLysSerProse 382
QY 554 -CCCGGATCCTCGGCTCCCTTGACGCTCTACCCATGGRGAGAGCCAGTTTGARKMTW 612
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
382 rAlaGlySerProGlyPro----- 388
QY 613 TACATGTTGGTGAAGAGAGAGACCGTGGATGGTACATGAATGAAGATGACCT---G 669
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
389 -----AlaArgGly-SerArgGlyTrpProTyArgValArgTrpProHisA 404
QY 670 CCCAGAGCGCTCGCTCCAGATCATCCGTGACCTTCGGCATATAATTCGCCAGTGAA 729
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
404 rArgArgProSerArgSerAlaArgArgAlaAlaGlyArgProAlaAlaAlaProA 424
QY 730 GAAATTACAGAGAGAGAGTT-----GGAGAACGCTGTCAGCAATTCCTCGAGAGA 783
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
424 rGSerAlaProGlySerValAlaAlaGlyArgProAlaArg-SerAlaArgSerArg- 442
QY 784 TATAACCGTTCACTGGGCTCTACTTGT-----CATCAATGC 819
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
443 -----ArgArgAlaProSerHisCysArgGlyProProGlyProGlyAsnSerCys 460
QY 820 CGTCAGAGACTATTGATACCAAAACAACTGACAGAAACCCAGACTCTCGGGCGTTCGA 879
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
461 -IleArgArg-----SerAlaThrAlaSerTrpProGlyAlaGlyAlaProAr 476
QY 880 GGCCAGTTCTGTGGCCCTCGCTTCCGAAACCGTTTATGTCGAAGAGGTCAGGATGCTCTG 939
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db 476 gCysSerProAlaHisAla---GluLeuValGlnAlaArgArg-----ValVa 492
Qy 940 CTGGATCCGACCTGCGCTTCCCGCTTTCGAGGATCTGCA-----ACTGAGTTTC 993
Db 492 lGluLeuArgLeuAlaMetValArgValValAlaSerAlaValGlyThrValLeuAr 512
Qy 994 TGCCGGCAGGAGATGACGCTGTCGACCTGGGGTCC 1030
Db 512 gLeuGluArgArgLeuAspGlyLeuHisArgGlyAla 524

RESULT 14
US-09-418-780A-1
; Sequence 1, Application US/09418780A
; Patent No. 6504009
; GENERAL INFORMATION:
; APPLICANT: Jones, Michael H.
; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATOR
; FILE REFERENCE: 06501-043001
; CURRENT APPLICATION NUMBER: US/09/418,780A
; CURRENT FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: PCT/JP98/01782
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: JAPAN 9/116402
; PRIOR FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-418-780A-1

Alignment Scores:
Pred. No.: 0.0359 Length: 947
Score: 103.50 Matches: 71
Percent Similarity: 33.44% Conservative: 37
Best Local Similarity: 21.98% Mismatches: 90
Query Match: 5.28% Indels: 125
DB: 4 Gaps: 16

US-10-046-935-2234 (1-1116) x US-09-418-780A-1 (1-947)

Qy 64 TATGTGAAGTTGATTTCATCGAAGACCTCG----- 93
Db 388 TyrIleLysThrAspIleThrGluThrThrGlyArgGluAsnThrAsnGluAlaSerSer 407
Qy 94 -----TCATCTCTGATGACAGTTGTGACAGCTTGTCTTGATAATTTTGCAACACG 147
Db 408 GluGlyAsnSerSerAspSerGluAspGluArgValLys----- 421
Qy 148 AGGCTGCAGTCAGTTCCGGAAGCTGTAGGACCCGACGAGGACCTCTGGACCT 207
Db 422 ArgLeuAlaLysLeuGlnGlnLeuLysAlaValHisGln-----Gln 436
Qy 208 CTCAGGGTGGCGATGAAGTTTCCAGCGCGAGTACCAGGGGAGCAACCAACAAAAACA 267
Db 437 LeuGlnValLeuSerGlnValProPheArgLys-----LeuAsnLysLysLys 452
Qy 268 GAG-----TCCCGCCAGCTTCAGAGAAATCTGTGACTGATTCCACTCCGAT----- 315
Db 453 GluLysSerLysLysLysLysLysLysLysLysValAsnAsnSerAsnGluAsnProArg 472
Qy 315 ----- 315
Db 473 LysMetCysGluGlnMetArgLeuLysGluLysSerLysArgAsnGlnProLysLysArg 492
Qy 316 -----TCAGAAGATGAAGTGA-----ATGAATTTT 342
Db 493 LysGlnGlnPheIleGlyLeuLysSerGluAspGluAspAsnAlaLysProMetAsnTyr 512
Qy 343 TTGGAGAAAGCGCTTAAATATAAGCAACAAACAAAGCAATCTTGCAAACTCATGCT 402
Db 513 AspGluLysArgGlnLeuSerLysLysLysLysLysLysLysLys----- 524

Qy 403 GAATTAGAAAGCTTCCTCGCTCGTTCCTCGTGAAGA-----CATCCCTCCCGAGGCTCC 456
Db 525 -----LeuProGlyAspLysLeuGlyArgValValHisIleIle----- 537
Qy 457 GACTCAATCAAGGAGACCGGACCGGTACATTCCCGGTGTGTTCCTCCAGGAGAAAC 516
Db 538 -----GlnSerArgGluProSer-----LeuSerAsnSerAsn 548
Qy 517 CCTGAACGG----- 525
Db 549 ProAspGluLeuGluLeuAspPheGluThrLeuLysAlaSerThrLeuArgGluLeuGlu 568
Qy 526 -----AGAGCTCGTCTCTTACCAAGGTCAAGGTCCCGGATCCTC 564
Db 569 LysTyrValSerAlaCysLeuArgLysArgProLeuLysProAlaLysLysIleMet 588
Qy 565 GGGTCCCTTACAGCTTACCCATGGRGAGAGCCAGTTTGAARKMTWGTACATCTTCGTG 624
Db 589 MetSerLysGluGluLeu-----HisSerGlnLysLysGlnGluLeuGlu 603
Qy 625 AGAAAGAGGAGACCGTGGATGCTACATGAATGAGATGACCTGCCAGGAGCCGTCGC 684
Db 604 LysArgLeuLeuAspValAsnAsnGlnLeuAsnSer-----ArgLysArgGln 619
Qy 685 TCCAGATCATCCGTGACCTTCCGCATATATTTGCCCGCAGTGGAGAAATTACA----- 738
Db 620 ThrLysSerAspLysThrGlnProSer-----LysAlaValGluAsnValSerArgLeu 637
Qy 739 GAGGAGAGTGGAGACGCTCGACCAATCTCGAGAGAGATATATAACGTTCACTG 798
Db 638 SerGluSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 657
Qy 799 GGCTCTACT 807
Db 658 SerSerSer 660

RESULT 15
US-09-392-714-23
; Sequence 23, Application US/09392714A
; Patent No. 6686147
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Gure, Ali O.
; APPLICANT: Williamson, Barbara
; APPLICANT: Chen, Yao-Tsung
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Cancer Associated Antigens and Uses
; FILE REFERENCE: L0461/7062
; CURRENT APPLICATION NUMBER: US/09/392,714A
; CURRENT FILING DATE: 1999-09-09
; EARLIER APPLICATION NUMBER: PCT/US98/14679
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-392-714-23

Alignment Scores:
Pred. No.: 0.0359 Length: 947
Score: 103.50 Matches: 71
Percent Similarity: 33.44% Conservative: 37
Best Local Similarity: 21.98% Mismatches: 90
Query Match: 5.28% Indels: 125
DB: 4 Gaps: 16

US-10-046-935-2234 (1-1116) x US-09-392-714-23 (1-947)

Qy 64 TATGTGAAGTTGATTTCATCGAAGACCTCG----- 93

B/6/K